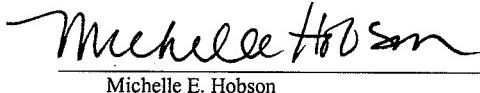


PATENT
USSN 09/844,662
Docket No. 8325-0012 (S12-US1)

I hereby certify that this correspondence is being submitted to the United States Patent and Trademark Office via EFS on **September 8, 2009**.


Michelle E. Hobson

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

E. Raschke *et al.*

Application No.: 09/844,662

Filed: April 27, 2001

For: METHODS FOR BINDING AN
EXOGENOUS MOLECULE TO
CELLULAR CHROMATIN

Examiner: Robert M. Kelly

Group Art Unit: 1633

Confirmation No.: 9004

BRIEF ON APPEAL UNDER 37 C.F.R. § 41.37

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This Appeal Brief is filed pursuant to 37 C.F.R. § 41.37 and is in response to the Final Office Action mailed on April 17, 2009 and Advisory Actions mailed June 25, 2009 and July 28, 2009. A Notice of Appeal was received in the USPTO on July 7, 2009, making an Appeal Brief due on or before September 7, 2009. As this Appeal Brief is filed on the first business day following the September 7, 2009 Labor Day holiday, it is timely filed.

REAL PARTY IN INTEREST

Sangamo BioSciences, Inc. is the assignee of record, based on an assignment from the inventors recorded on October 16, 2001 at Reel 012068, Frame 0811. Thus, Sangamo BioSciences, Inc. are the real parties in interest.

RELATED APPEALS AND INTERFERENCES

Appellants are not aware of any appeals or interferences in this series.

However, this particular Examiner has required that Appellants submit a list of all pending appeals in cases assigned to Sangamo BioSciences (the Real Party in Interest) in at least 2 other cases on Appeal (U.S. Serial No. 10/986,583 and 11/148,794). As no decisions in these cases have yet been issued, Appellants are awaiting the Board's guidance regarding the requirement that all cases assigned to Sangamo and containing the term "zinc finger protein" in the claims are "related" as defined in 37 C.F.R. § 41.37(c)(ii).

Nevertheless, as will be required by the Examiner, Appellants note that they are aware of the following additional Appeals in cases in which Sangamo BioSciences is the Real Party in Interest:

U.S. Serial No. 09/825,242, filed April 21, 2001, now U.S. Patent No. 7,177,766

U.S. Serial No. 09/424,482, filed February 29, 2000, now abandoned

U.S. Serial No. 09/844,508, filed April 27, 2001, now U.S. Patent No. 7,001,768

U.S. Serial No. 10/084,826, filed October 24, 2001, now abandoned

U.S. Serial No. 09/636,243, filed August 10, 2000

U.S. Serial No. 10/470,180, filed December 22, 2003, now U.S. Patent No. 7,262,054

U.S. Serial No. 10/337,216, filed January 6, 2003, now U.S. Patent No. 7,491,531

U.S. Serial No. 09/844,501, filed April 27, 2001, now U.S. Patent No. 7,217,509

U.S. Serial No. 10/083,682, filed October 24, 2001

U.S. Serial No. 09/996,484, filed November 28, 2001

U.S. Serial No. 10/397,930, filed March 25, 2003, Board Decision reversing the Examiner's rejections decided on August 29, 2009

PATENT
USSN 09/844,662
Docket No. 8325-0012 (S12-US1)

U.S. Serial No. 10/400,017, filed March 25, 2003

U.S. Serial No. 10/474,282, filed April 26, 2004

U.S. Serial No. 10/418,552, filed April 17, 2003

U.S. Serial No. 10/198,677, filed July 17, 2002

U.S. Serial No. 10/470,065, filed February 5, 2004

U.S. Serial No. 10/651,761, filed August 29, 2003, now U.S. Patent No.

7,361,635

U.S. Serial No. 10/986,583, filed November 12, 2004

U.S. Serial No. 11/304,981, filed December 15, 2005

U.S. Serial No. 11/101,095, filed April 7, 2005, now U.S. Patent No. 7,534,775

U.S. Serial No. 10/656,531, filed September 5, 2003

STATUS OF CLAIMS

Pending: Claims 57, 68-71, 87-91 and 96-102

Canceled: Claims 1 to 56, 58-67, 72-90, 92-95

Withdrawn: Claims 91, 93, and 96-102

Rejected: Claims 57 and 68-71

Appealed: Claims 57 and 68-71

STATUS OF AMENDMENTS

The amendments made in the paper responsive to the Final Office Action were not entered for purposes of Appeal by the Examiner. (Advisory Actions, Box 7). In particular, Appellants attempted to amend the claims to remove the term "non-naturally occurring" in line 2 (modifying zinc finger protein).

Thus, the claims on appeal are as shown in the attached Claims Appendix.

SUMMARY OF THE CLAIMED SUBJECT MATTER

Independent claim 57 is drawn to cell comprising a complex (page 7, lines 10-14) between (i) a non-naturally occurring zinc finger protein (page 17, lines 4-20) comprising 3 or more zinc finger domains, wherein the zinc finger domains comprise a

non-naturally occurring recognition helix (page 27 (Table 1)) and (ii) chromosomal cellular chromatin (page 10, lines 10-14). Furthermore, the zinc finger protein is bound to a target site in a region of the cellular chromatin that is sensitive to digestion with DNaseI (page 23, lines 10-25).

Claim 68 depends from claim 57 and further specifies that the zinc finger protein is encoded by a nucleic acid introduced into the cell (page 7, lines 3-4; Examples).

Claim 69 depends from claim 57 and further specifies that the cell is a plant cell (page 6, line 11; page 7, line 9; page 12, lines 29-30).

Claim 70 depends from claim 57 and further specifies that the cell is an animal cell (page 6, line 11; page 7, line 9; page 12, lines 29-30).

Claim 71 depends from claim 57 and further specifies that the cell is a human cell (page 6, line 12; page 7, line 9; page 12, lines 29-30).

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

A. Whether claims 57 and 68-71 are unpatentable under 35 U.S.C. § 112, 1st paragraph (written description) for containing new matter not disclosed in the as-filed specification.

B. Whether claims 57, 68, 70 and 71 are unpatentable 35 U.S.C. § 102(e) as anticipated by U.S. Patent No. 6,013,453 (hereinafter “Choo”).

C. Whether claims 57 and 68-71 are unpatentable under 35 U.S.C. § 103(a) as obvious over Choo in view of WO 00/9837755 (hereinafter “Dangl”).

ARGUMENTS

A. The claims on appeal do not contain new matter

Claims 57 and 68-71 remain rejected under 35 U.S.C. § 112, 1st paragraph allegedly containing new matter not described in the originally filed specification, namely a non-naturally occurring zinc finger protein comprising at least one zinc finger domain with a non-naturally occurring recognition helix. (Final Office Action, pages 11-13). It was alleged that the term “non-naturally occurring” as used in both contexts was unclear, undefined and not described by the as-filed specification. *Id.*

In a sincere effort to advance prosecution and clarify that the zinc finger proteins as a whole are non-naturally occurring by virtue of the non-naturally occurring recognition helices of the individual zinc fingers, after final, Appellants amended claim 57 (line 2) to remove the term “non-naturally occurring” before the recitation “zinc finger protein” in line 2. However, the Examiner did not enter the amendment, asserting that it raised issues of new matter and would require a new search. (Advisory Actions, page 2).

Furthermore, in response to Appellants’ clarification that the claims are not directed to zinc finger proteins with at least 3 fingers in which one of the fingers “recognizes a generic non-naturally finger is bound to cellular chromatin in a region sensitive to digestion with DNaseI” (see, Final Office Action, page 12), the Examiner asserted that the term non-naturally occurring is not described (Advisory Action, dated June 25, 2009, date page 2):

It would appear that Applicant did not contemplate in their specification that design/modification could yield anything but non-naturally occurring zinc fingers/sequences. However, Applicant, not being cognizant of such, specifically means that Applicant did not possess the breadth of what Applicant is claiming. If Applicant wants to limit the claim to those that are designed or selected, it would appear that Applicant should claim sequences that are designed or selected. However, attempting to subvert the plain meaning of words, to remove their black-and-white meaning, is antithetical to plain English, and hence, is not allowed. While Applicant may be their own text, Applicant has not provided a definition, and if such definition removed the plain of the words, then it would be improper anyway. Applicant argues that there exists much art to demonstrate possession of selection/designing sequences. Such is not persuasive. It has nothing to do with the rejections.

It is well settled that the proscription against the introduction of new matter in a patent application (35 U.S.C. 132 and 251) serves to prevent an applicant from adding information that goes beyond the subject matter originally filed. See, e.g., *In re Rasmussen*, 650 F.2d 1212, 1214, 211 USPQ 323, 326 (CCPA 1981) and MPEP § 2163.06. Furthermore, literal support is not required, when, in fact, M.P.E.P. § 2163.02 specifically indicates the reverse, namely:

The subject matter of the claim need not be described literally (*i.e.*, using the same terms or *in haec verba*) in order for the disclosure to satisfy the description requirement.

Thus, the written description requirement is satisfied if the specification reasonably conveys possession of the invention to one skilled in the art. *See, e.g., In re Lukach*, 169 USPQ 795, 796 (CCPA 1971). Moreover, the burden is on the Examiner to provide evidence as to why a skilled artisan would not have recognized that the applicant was in possession of claimed invention at the time of filing. *Vas-Cath, Inc. v. Mahurkar*, 19 USPQ2d 1111 (Fed. Cir. 1991); *In re Wertheim*, 191 USPQ 90 (CCPA 1976).

In the case on appeal, the skilled artisan would have no doubt as to the meaning of the recitations of a “non-naturally occurring zinc finger protein” of at least 3 fingers in which the 3 fingers each “comprise a non-naturally occurring recognition helix” and that Appellants were in possession of such non-naturally occurring zinc finger proteins at the time of filing.

Naturally occurring and non-naturally occurring recognition helices of zinc fingers are described in detail in the as-filed specification and were well-known to the skilled artisan at the time of filing. *See*, page 5, lines 14-25 (paragraph [0019] pf published application); page 6, lines 20-31 (paragraph [0025] of the published application); page 17, lines 4-20 (paragraph [0070] of published application); Example 8 (paragraphs [0110], [0111] and [0114] of published application); Table 1 which shows exemplary non-naturally occurring zinc finger recognition domains; Example 15 (paragraph [0158] of published application), emphasis added):

In another embodiment, an accessible region is identified within a region of interest and a ZFP target site is located within the accessible region. A ZFP that binds to the target site is designed. The designed ZFP can be introduced into the cell, or a nucleic acid encoding the designed ZFP can be designed and the designed nucleic acid can be introduced into the cell, where it will express the designed ZFP.
Methods for the design and/or selection of ZFPs that bind specific sequences are disclosed in U.S. Pat. No. 5,789,538; U.S. Pat. No. 6,007,408; U.S. Pat. No. 6,013,453; PCT WO 95/19431; PCT WO 98/54311 co-owned PCT/US00/00388

and references cited therein; co-owned U.S. patent application Ser. No. 09/444,241, filed Nov. 19, 1999; and co-owned U.S. patent application Ser. No. 09/535,088, filed Mar. 23, 2000. Methods for selection include, but are not limited to, phage display and in vivo selection.

In methods comprising introduction of an exogenous molecule into a cell and testing for binding of the exogenous molecule to a binding site, a ZFP that binds to a target site, located within an accessible region, is designed. The designed ZFP can be introduced into the cell, or a nucleic acid encoding the designed ZFP can be designed and the designed nucleic acid can be introduced into the cell, where it will express the designed ZFP. Methods for the design and/or selection of ZFPs that bind specific sequences are disclosed in U.S. Pat. No. 5,789,538; U.S. Pat. No. 6,007,408; U.S. Pat. No. 6,013,453; PCT WO 95/19431; PCT WO 98/54311 co-owned PCT/US00/00388 and references cited therein; co-owned U.S. patent application Ser. No. 09/444,241, filed Nov. 19, 1999; and co-owned U.S. patent application Ser. No. 09/535,088, filed Mar. 23, 2000. Methods for selection include, but are not limited to, phage display and in vivo selection.

In a preferred embodiment, an exogenous molecule is a zinc finger DNA-binding protein (ZFP). Certain ZFPs, their properties and their binding sequences are known in the art, as described supra. Furthermore, it is possible, for any particular nucleotide sequence, to design and/or select one or more ZFPs capable of binding to that sequence and to characterize the affinity and specificity of binding. See, for example, U.S. Pat. No. 5,789,538; U.S. Pat. No. 6,007,408; U.S. Pat. No. 6,013,453; PCT WO 95/19431; PCT WO 98/54311 co-owned PCT/US00/00388 and references cited therein; co-owned U.S. patent application Ser. No. 09/444,241, filed Nov. 19, 1999; and co-owned U.S. patent application Ser. No. 09/535,088, filed Mar. 23, 2000. Certain sequences, such as those that are G-rich, are preferred as ZFP binding sites. Since a three-finger ZFP generally binds to a 9- or 10-nucleotide target site, in a preferred embodiment, an accessible region, present within a region of interest in cellular chromatin, is searched for one or more G-rich sequences of 9-10 nucleotides and, for each sequence so detected, a ZFP can be designed to bind those sequences. In addition, two three finger modules can

be joined, via an appropriate linker domain, to form a six-finger protein capable of recognizing an 18-20 nucleotide target site. See, for example, PCT/US99/04441.

Plasmids were constructed to encode transcriptional effector proteins containing zinc finger domains designed to recognize target sites surrounding the transcriptional initiation site of the human vascular endothelial growth factor (VEGF) gene; i.e. within the +1 accessible region described in Example 7. The target site has the sequence 5'-GGGGAGGATCGCGGAGGCTT-3' (SEQ ID NO: 1), where the underlined T residue represents the major transcriptional startsite for the VEGF gene. A binding domain containing six zinc fingers, named VEGF 3a/1, was designed to bind to this 20-nucleotide target sequence. A three-finger zinc finger domain, VEGF 1 was designed to bind to the upstream 10-nucleotides of this target site having the sequence 5'-GGGGAGGATC-3' (SEQ ID NO: 2). A control six-finger domain, GATA 15.5, which was designed to bind the sequence 5'-GAGTGTGTGAACTGCGGGGCAA-3' (SEQ ID NO: 3), was also used. These zinc finger domains were encoded as fusion proteins in the NVF vector, as described below.

The zinc finger domains were constructed in a SP1 backbone. The sequences of the recognition helices, from position -1 to position +6, of VEGF 3a/1, VEGF 1 and GATA 15.5 are shown in Table 1. ...

The zinc finger domains contained **designed recognition helices**, as shown in Table 1, in a SP1 backbone.

An engineered fusion protein was designed to recognize a unique 9 base pair sequence in the DNase I hypersensitive region at -2 kb. This protein (BOS 3) comprised a nuclear localization sequence, a zinc finger binding domain, a KRAB repression domain and a FLAG epitope. The zinc finger binding domain was targeted to the sequence GGGGAGGAG, (SEQ ID NO: 27) which is complementary to the sequence CTCCTCCCC (SEQ ID NO: 28) in the coding strand. Zinc finger sequences (for amino acids -1 through +6 of the recognition helices) were RSDNLTR (SEQ ID NO: 29), RSDNLTR (SEQ ID NO: 30) and RSDALT (SEQ ID NO: 31). Construction of a plasmid encoding the fusion protein and determination

of the binding affinity of the zinc finger binding domain for its target sequence were performed according to methods disclosed in co-owned PCT WO 00/41566 and WO 00/42219. The dissociation constant (Kd) was determined to be 3.5 pM.

Thus, the recitation “non-naturally occurring recognition helix” is clearly referring to a recognition helix that does not occur in nature, but has been designed and/or selected to bind to a particular target site using well known methods. The use of this term is not antithetical to its plain meaning, but, rather, completely in keeping with the art-recognized and specification-detailed definition.

Indeed, the well-understood definition of the term “non-naturally occurring” was recently addressed by the Board of Patent Appeals and Interferences, where the Board confirmed that even when the phrase “naturally occurring” does not appear verbatim in the specification, it would clearly be understood by the skilled artisan to mean something that exists or is found in nature. See, *Ex parte Dewis* (Appeal 2007-1610, decided September 4, 2007). Plainly, the skilled artisan would have no doubt as to the scope of the term “non-naturally occurring,” with respect to zinc finger proteins as a whole and to their individual recognition helix regions, as referring to zinc finger proteins whose recognition helices are not found in nature (i.e., are designed and/or selected to bind to a particular target site).

Appellants also strongly traverse the assertion that the vast amount art regarding zinc finger proteins that include zinc fingers with non-naturally occurring recognition helices is not relevant to the rejection. In fact, it is axiomatic that for determining compliance with 35 U.S.C. § 112, 1st paragraph, written description, the disclosure must be read in light of the knowledge possessed by the skilled artisan at the time of filing, for example as established by reference to patents and publications available to the public prior to the filing date of the application. See, e.g., *In re Lange*, 209 USPQ 288 (CCPA 1981).

In the case on appeal, there are a myriad of references cited in the specification regarding zinc finger proteins containing non-naturally occurring recognition helices and how to make and use these proteins. See, e.g., col. 2, lines 64-67; col. 1, lines 35-43; col.

5, lines 7-24) of U.S. Patent No. 5,789,538, cited on page 5 of the as-filed specification and reference AA-1 of IDS submitted May 3, 2002, and Evidence Appendix (1) of this Brief, emphasis added):

Naturally occurring zinc fingers, referred to herein as "wild type" zinc fingers occur in regulatory fractions that include differentiation and growth signals...

In the present invention, the polypeptide encompassing the zinc fingers 1) differs in sequence--at one or more base-contacting amino acid residues--from a known wild type zinc finger protein or has a base sequence specificity distinct from that of a known wild type zinc finger protein and 2) binds to a targeted polynucleotide with high affinity (binds DNA with high affinity, i.e., with a dissociation constant of less than 1 nanomolar (nM)).

Preferably, finger 1 would have mutations at one or more of the base-contacting positions; i.e. finger 1 would have the same polypeptide sequence as a wild type zinc finger except that at least one of the four amino residues at the primary positions would differ. Similarly, to design a three-finger zinc protein that would bind to a 10 base sequence ABCXXXXH₂IJ (wherein each "X" is base that would be specified in a particular application), fingers 1 and 3 have the same sequence as the wild type zinc fingers which bind GH₂IJ and ABCD, respectively, while finger 2 would have residues at one or more base-coating positions which differ from those in a wild type finger. The present invention encompasses multifingered proteins in which more than one finger differs from a wild type zinc finger. It also includes multifingered protein in which the amino acid sequence in all the fingers have been changed.

See, also, col. 3, lines 28-40; col. 17, lines 46-54 of U.S. Patent No. 6,013,453, cited on page 5 of the specification and Ref. AC-1 of IDS filed May 3, 2002, relied on by the Examiner in the anticipation and obviousness rejections and Evidence Appendix (2) this Brief:

In summary, it is known that Zf protein motifs are widespread in DNA binding proteins and that binding is via three key amino acids, each one contacting a single base pair in the target DNA sequence. Motifs are modular and may be linked together to form a set of fingers which recognise a contiguous DNA sequence (e.g. a three fingered protein will

recognise a 9mer etc). The key residues involved in DNA binding have been identified through sequence data and from structural information. Directed and random mutagenesis has confirmed the role of these amino acids in determining specificity and affinity. Phage display has been used to screen for new binding specificities of random mutants of fingers.

The amino acid sequence biases observed within a family of functionally equivalent zinc fingers indicate that, of the α -helical positions randomised in this study, only three primary (-1, +3 and +6) and one auxiliary (+2) positions are involved in recognition of DNA. Moreover, a limited set of amino acids are to be found at those positions, and it is presumed that these make contacts to bases. The indications therefore are that a code can be derived to describe zinc finger-DNA interactions.

Thus, the term “non-naturally occurring” as used to refer to the recognition helix is clearly used in the claims in both an art-recognized and specification-defined manner.

For the reasons of record and reiterated herein, the claimed subject matter is fully described in the as-filed specification and no new matter has been added and Appellants were in possession of the claimed complexes at the time of filing. Accordingly, the rejection cannot be sustained.

B. Claims 57, 68, 70 and 71 are not anticipated by Choo

Claims 57, 68, 70 and 71 remain rejected under 35 U.S.C. § 102(e) as allegedly anticipated by U.S. Patent No. 6,013,453 (hereinafter “Choo”), which was alleged to disclose the making of a mutant 3-fingered zinc finger protein that binds to the “coding sequence for a specific ras mutation” in human cells. (Final Office Action, page 16). It was also alleged that “absent reason to believe otherwise, this site occurs within the broad definition of a general region which is in some way sensitive to digestion with DNaseI.”

Id.

In response to Appellants arguments that Choo does not necessarily and inevitably teach that their zinc finger proteins form a complex with a region of chromosomal cellular chromatin that sensitive to DNaseI digestion, it was asserted that “Choo is not required to state that they did it, but only teach that it can be done.” (Advisory Action, dated June 25, 2009, page 2).

However, in order to support an anticipation rejection, the Office must show that every limitation of the claim at issue must appear identically in a single reference. *In re Bond*, 910 F.2d 831, 832, 15 USPQ2d 1566, 1567 (Fed. Cir. 1990). Furthermore inherent anticipation cannot be established by probabilities or possibilities (see, *Continental Can Co. USA , Inc. v. Monsanto Co.*, 20 USPQ2d 1746, 1749 (Fed. Cir. 1991):

Inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient.

Thus, in the instant case, the rejection cannot be supported when Choo does not teach what is claimed.

In the case on appeal, the claims require that the non-naturally occurring zinc finger protein of the complex be bound to a site in chromosomal cellular chromatin. Furthermore, the zinc finger protein is necessarily and inevitably bound to DNaseI hypersensitive sites in chromosomal cellular chromatin.

As a threshold matter, it remains the case that Choo fails to teach non-naturally occurring proteins that are necessarily and inevitably bound to a chromosomal cellular chromatin. As clearly defined in the specification, chromosomal cellular chromatin is endogenous chromatin (see, page 1, lines 10-12; page 10, lines 7 to 17 of the as-filed specification):

The present disclosure is in the field of gene regulation, specifically, regulation of an endogenous gene in a cell and methods of regulating an endogenous gene through binding of an exogenous molecule.

Chromatin is the nucleoprotein structure comprising the cellular genome. Cellular chromatin comprises nucleic acid, primarily DNA, and protein, including histones and non-histone chromosomal proteins.

A chromosome, as is known to one of skill in the art, is a chromatin complex comprising all or a portion of the genome of a cell. The genome of a cell is often characterized by its karyotype, which is the collection of all the chromosomes that comprise the genome of the cell. The genome of a cell can comprise one or more chromosomes.

An episome is a replicating nucleic acid, nucleoprotein complex or other structure comprising a nucleic acid that is not part of the chromosomal karyotype of a cell. Examples of episomes include plasmids and certain viral genomes.

Simply put, the claims require that the complex be between chromosomal cellular chromatin, which is necessarily endogenous, and a non-naturally occurring zinc finger protein.

By contrast, Choo fails to teach binding of a zinc finger protein to chromosomal cellular chromatin. Rather, Choo's Bcr-Abl site is synthesized and then inserted into the genome. Thus, Choo's protein is not binding to endogenous chromosomal cellular chromatin, but to a sequence inserted into the genome.

The Examiner also errs in asserting that the zinc finger protein of Choo is necessarily and inevitably bound to an accessible region in cellular chromatin. Indeed, Choo teaches that their inserted site is not packaged within chromatin, as the claimed site within chromosomal cellular chromatin would necessarily be (see, Choo, col. 28, lines 39-47, emphasis added):

The selective stimulation of transcription indicates convincingly that highly site-specific DNA-binding can occur *in vivo*. However, while transient transfections assay binding plasmid DNA, the true target site for this and most other DNA-binding proteins is genomic DNA. This might well present significant problems, not least since this DNA is physically separated from the cytosol by the nuclear membrane, but also since it may be packaged within chromatin.

Thus, Choo clearly states that endogenous cellular chromatin as recited in the claims is different from their randomly integrated sequences (or their episomal reporter plasmids) that are not in their normal chromatin environment. In other words, a sequence randomly integrated into a cell's genome is not chromosomal cellular chromatin, as claimed.

Therefore, Choo does not describe or demonstrate complexes of zinc finger binding proteins with cellular chromosomal chromatin, as claimed. As such, Choo cannot anticipate any of the pending claims and the rejection cannot be sustained.

C. Claims 57 and 68-71 are not obvious over the cited references

Claims 57 and 68-71 also remain rejected under 35 U.S.C. § 103(a) as allegedly obvious over Choo in view of WO 00/9837755 (hereinafter “Dangl”). (Final Office Action, pages 16-17). Choo was cited as above and Dangl was cited for teaching that zinc finger proteins function in plant cells. *Id.*

In regard to Appellants’ arguments that showing binding of a zinc finger protein to randomly integrated is not the same as (or predictive of) binding to endogenous chromosomal cellular chromatin, it was asserted that Choo’s randomly integrated sequence is a chromosomal sequence as claimed. (Advisory Action, dated June 25, 2009, pages 2-3).

Appellants first note that the obviousness rejection is improper because the cited references, alone or in combination, fail to teach the elements of the claims. In fact, in order to establish obviousness of a claimed invention, all the features of the claims must be taught or suggested by the prior art. *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974).

As noted above and throughout prosecution, Choo does not teach complexes in which a non-naturally occurring zinc finger protein is bound to a site in chromosomal cellular chromatin. Instead, Choo’s zinc finger proteins bind to randomly inserted sequences and these randomly integrated sequences are not cellular chromatin. Likewise, Dangl fails to teach or suggest complexes as claimed. Therefore, on this basis alone, the rejection cannot stand.

Appellants also traverse the assertion that additional evidence of record establishes that Choo’s zinc finger proteins are not bound to cellular chromatin as claimed. (Advisory Action dated June 25, 2009, pages 2-3). Indeed, Appellants have provided ample evidence establishing that binding of zinc finger proteins to a sequence randomly inserted into the cell is entirely unlike binding of zinc finger proteins to chromosomal cellular chromatin as claimed. In particular, as disclosed in Beerli et al. (2000) *Proc. Natl. Acad. Sci. USA* 97:1495-1500, page 1465, left column, reference AH-

1 of IDS submitted May 3, 2002¹ and Evidence Appendix (3) of this Brief, well after Choo's work, the skilled artisans working in the field of zinc finger proteins clearly indicated that they regarded modulating endogenous gene expression with a non-naturally occurring zinc finger protein as an unmet challenge:

While our early experiments have focused on the regulation of genes transiently introduced into cells, we realized that the willful and specific regulation of endogenous genes with designed transcription factors has remained an unmet challenge in biology.

In a separate discussion article, a co-author of this Beerli paper (published in 2000) also discussed the complete lack of predictability of using ZFPs to bind to endogenous genes based on studies such as Choo's (which used introduced binding sites):

"This is the first time we've been able to show that these designed transcription factors work on real genes and real chromosomes, not genes of binding sites that have been introduced into cells,"....²

Thus, the issue is not whether Choo is enabling for what it describes, namely binding of a zinc finger protein to an exogenous sequence randomly inserted into the genome. The issue is whether Choo teaches binding of a zinc finger protein to endogenous chromosomal cellular chromatin as claimed. Clearly, in view of the fact that the skilled artisan working in this field is on the record as stating that binding of engineered zinc finger proteins to chromosomal cellular chromatin was not predictable based on Choo's disclosure, Choo does not teach the claimed elements or in any way suggest their predictability.

Therefore, there is no combination of Choo and Dangl that render any of the pending claims obvious over these references and the rejection must be withdrawn.

¹ Appellants note that the Examiner's assertion on page 3 of the Advisory Action dated June 25, 2009 that this reference was not submitted in an IDS is incorrect (see, also Evidence Appendix).

² See, Borman "DNA-Binding Proteins Turn Genes On and Off." February 21, 2000, C&CEN (Evidence Appendix (4)).

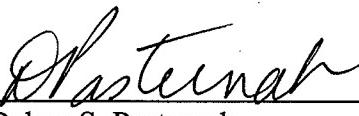
PATENT
USSN 09/844,662
Docket No. 8325-0012 (S12-US1)

CONCLUSION

For the reasons set forth herein, allowance of the claims under consideration is requested.

Respectfully submitted,

Date: September 8, 2009

By: 
Dahna S. Pasternak
Registration No. 41,411
Attorney for Appellants

ROBINS & PASTERNAK LLP
1731 Embarcadero Road, Suite 230
Palo Alto, CA 94303
Telephone: (650) 493-3400
Facsimile: (650) 493-3440

CLAIMS APPENDIX

- 57.** A cell comprising a complex between
a non-naturally occurring zinc finger protein comprising 3 or more zinc finger
domains, wherein the zinc finger domains comprise a non-naturally occurring recognition
helix and
chromosomal cellular chromatin;
wherein the zinc finger protein is bound to a target site in a region of the cellular
chromatin that is sensitive to digestion with DNaseI.
- 68.** The cell of claim 57, wherein the zinc finger protein is encoded by a nucleic
acid introduced into the cell.
- 69.** The cell of claim 57, wherein the cell is a plant cell.
- 70.** The cell of claim 57, wherein the cell is an animal cell.
- 71.** The cell of claim 57, wherein the cell is a human cell.

EVIDENCE APPENDIX

The following documents are submitted with this Brief:

- (1) U.S. Patent No. 5,789,538, cited reference AA-1 of IDS submitted May 3, 2002, indicated considered by the Office on July 19, 2004;
- (2) U.S. Patent No. 6,013,453, cited as reference AC-1 of IDS submitted on May 3, 2002, indicated considered by the Office on July 19, 2004;
- (3) Beerli et al. (2000) *Proc. Nat'l. Acad. Sci. USA* 97:1495-1496, cited as reference AH-1 of IDS submitted on May 3, 2002, indicated considered by the Office on July 19, 2004; and
- (4) Borman et al. (2000) "DNA-Binding Proteins Turn Genes On and Off." C&CEN, submitted with the Response After Final dated June 10, 2009.

RELATED APPEALS AND INTERFERENCES

As noted above, Appellants attach the following Decisions on Appeal:

- (1) Decision on Appeal mailed September 27, 2006 in U.S. Serial No. 10/222,614; and
- (2) Decisions on Appeal mailed September 26, 2008 and March 26, 2009 in U.S. Serial No. 10/984,304;
- (3) Decision on Appeal mailed June 22, 2006 in U.S. Serial No. 09/825,242;
- (4) Decision on Appeal decided August 23, 2007 in U.S. Serial No. 09/424,482;
- (5) Decision on Appeal mailed May 30, 2007 in U.S. Serial No. 09/636,243;
- (6) Decision on Appeal mailed June 30, 2008 in U.S. Serial No. 10/337,216;
- (7) Decision on Appeal mailed December 21, 2006 in U.S. Serial No. 09/844,501;
- (8) Decision on Appeal mailed September 15, 2008 in U.S. Serial No. 10/083,682;
- (9) Decision on Appeal decided November 16, 2007 in U.S. Serial No. 10/651,761;
- (10) Decision on Appeal mailed September 30, 2008 in U.S. Serial No. 11/101,095;
- (11) Decision on Appeal decided August 28, 2009 in U.S. Serial No. 10/397,930



US005789538A

United States Patent [19]

Rebar et al.

[11] Patent Number: 5,789,538
[45] Date of Patent: Aug. 4, 1998

[54] ZINC FINGER PROTEINS WITH HIGH AFFINITY NEW DNA BINDING SPECIFICITIES

[75] Inventors: Edward J. Rebar, Cambridge; Carl O. Pabo, Newton, both of Mass.

[73] Assignee: Massachusetts Institute of Technology, Cambridge, Mass.

[21] Appl. No.: 850,250

[22] Filed: Apr. 18, 1997

Related U.S. Application Data

[63] Continuation of Ser. No. 383,056, Feb. 3, 1995, abandoned.

[51] Int. Cl.⁶ C07K 14/00; C07K 14/435; C07K 14/46; C07K 14/82

[52] U.S. Cl. 530/324; 530/300; 530/350; 530/358; 530/400

[58] Field of Search 530/300, 324, 530/350, 358, 400; 435/69.1, 172.3

[56] References Cited

U.S. PATENT DOCUMENTS

5,096,815 3/1992 Ladner et al. 435/69.1
5,223,409 6/1993 Ladner et al. 435/69.7

FOREIGN PATENT DOCUMENTS

WO95/19431 7/1995 WIPO

OTHER PUBLICATIONS

International Search Report for (AL) WO95/19431.

Thiesen, H. J. and Bach, C., "Determination of DNA Binding Specificities of Mutated Zinc Finger Domains," *Federation of European Biochemical Societies*, 283(1):23-26 (1991).

Nardelli, J., et al., "Zinc Finger-DNA Recognition: Analysis of Base Specificity by Site-Directed Mutagenesis," *Nucleic Acids Research*, 20(16):4137-4144 (1992).

Desjarlais, J. R. and Berg, J. M., "Redesigning the DNA-Binding Specificity of a Zinc Finger Protein: A Data Base-Guided Approach," *PROTEINS: Structure, Function and Genetics*, 12:101-104 (1992).

Desjarlais, J. R. and Berg, J. M., "Toward Rules Relating Zinc Finger Protein Sequences and DNA Binding Site Preferences," *Proc. Natl. Acad. Sci. USA*, 89:7345-7349 (Aug. 1992).

Desjarlais, J. R. and Berg, J. M., "Use of a Zinc-Finger Consensus Sequence Framework and Specificity Rules to Design Specific DNA Binding Proteins," *Proc. Natl. Acad. Sci. USA*, 90:2256-2260 (Mar. 1993).

Jamieson, A. C., et al., "In Vitro Selection of Zinc Fingers with Altered DNA-Binding Specificity," *Biochemistry*, 33:5689-5695 (1994).

Choo, Y. and Klug A., "Toward a Code for the Interactions of Zinc Fingers with DNA: Selection of Randomized Fingers Displayed on Phage," *Proc. Natl. Acad. Sci. USA*, 91:11163-11167 (1994).

Choo Y. and Klug A., "Selection of DNA Binding Sites for Zinc Fingers Using Rationally Randomized DNA Reveals Coded Interactions," *Proc. Natl. Acad. Sci. USA*, 91:11168-11172 (1994).

Wu, H., et al., "Building Zinc Fingers by Selection: Toward a Therapeutic Application," *Proc. Natl. Acad. Sci. USA*, 92:344-348 (1995).

Choo, Y., et al., "In Vivo Repression by a Site-Specific DNA-Binding Protein Designed Against an Oncogenic Sequence," *Nature*, 372:642-645 (1994).

Rice, W. G. et al. *Science* 270:1194-1197 (1995).

Taylor, W. E. et al. *Biochemistry* 34:3222-3230 (1995).

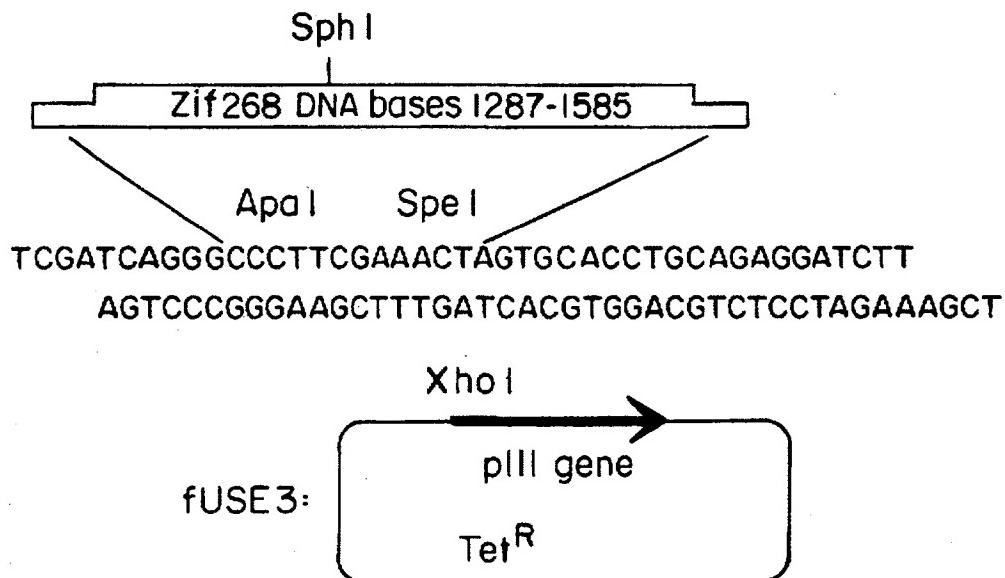
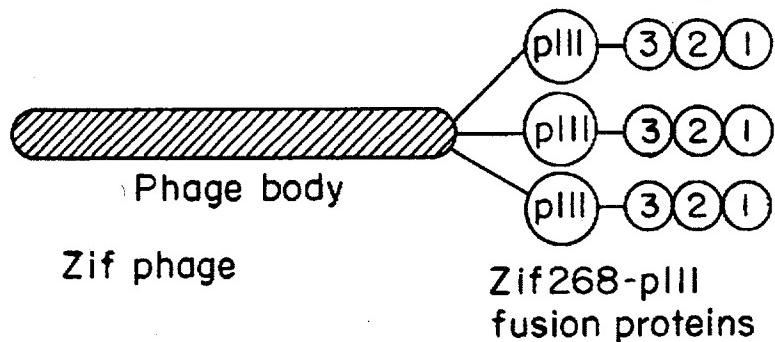
Primary Examiner—Dian C. Jacobson

Attorney, Agent, or Firm—Hamilton, Brook, Smith & Reynolds, P.C.

[57] ABSTRACT

Described is a polypeptide comprising one or more zinc fingers. The polypeptide binds to new polynucleotide subsites with high affinity and consequently has a binding specificity that differs from wild type zinc finger proteins. The binding occurs through contacts between certain amino acid residues of the zinc fingers and the nucleic acids of the subsites. The polypeptide sequence of at least one zinc finger differs from wild type zinc fingers, and the difference involves at least one amino acid residue that contacts the bases of the polynucleotide during binding.

8 Claims, 3 Drawing Sheets

**FIG. 1****FIG. 2**

GACC
GCAC
CCTG

CTGA[GCGTGG]_AGTGATCGATC-biotin
GACT[CGCACCx xx]TCAC

Fig. 3

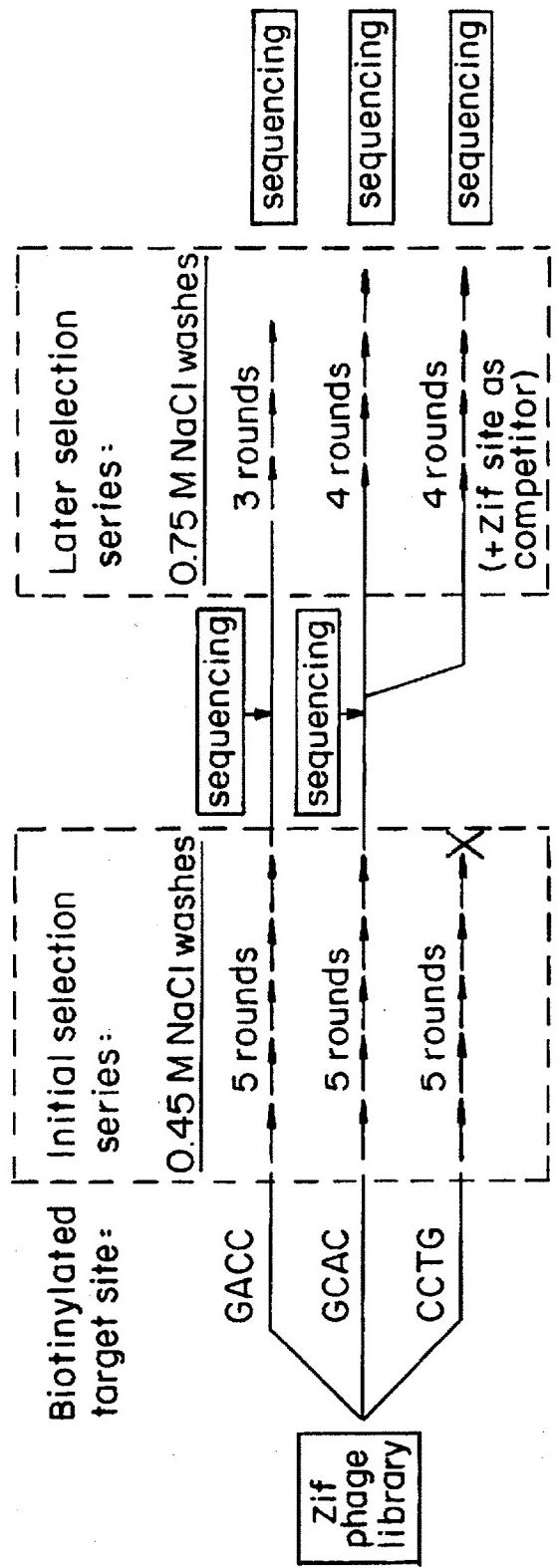


FIG. 4

1

ZINC FINGER PROTEINS WITH HIGH AFFINITY NEW DNA BINDING SPECIFICITIES

RELATED APPLICATION

This application is a continuation of application Ser. No. 08/383,056 filed on Feb. 3, 1995, now abandoned which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

There are many diseases which are caused by the overexpression of certain genes. For example, replication of viruses such as human immunodeficiency virus (HIV) are facilitated by the expression of certain viral genes by host cells. Many cancers are caused by the overexpression of certain genes, referred to as oncogenes. Such viral infection could be treated and the growth of such tumors could be prevented if the expression of the genes could be suppressed. Proteins which bind to the promoter sites of these genes would prevent their expression. However, there exists a need for proteins which can bind to the specific polynucleotide sequences of these promoter regions and which can be targeted to the cells which are overexpressing the undesired proteins.

SUMMARY OF THE INVENTION

The present invention is a polypeptide having three or more zinc fingers. The polypeptide binds specifically to a unique DNA site with high affinity. Zinc fingers typically bind to a span of three to four polynucleotide base pairs, referred to as a "subsite".

Another embodiment is a deoxyribonucleic acid that comprises a gene which encodes for the polypeptide of the present invention.

In the present invention, the polypeptide encompassing the zinc fingers 1) differs in sequence—at one or more base-contacting amino acid residues—from a known wild type zinc finger protein or has a base sequence specificity distinct from that of a known wild type zinc finger protein and 2) binds to a targeted polynucleotide with high affinity (binds DNA with high affinity, i.e., with a dissociation constant of less than 1 nanomolar (nM)). In one embodiment, a polypeptide of the present invention differs in the identity of its base-contacting amino acid residues from a corresponding wild type zinc finger protein, such as Zif268. In another embodiment, the polypeptide has a DNA binding specificity different from that of any known zinc finger protein. The base sequence specificity of a protein can be determined by selecting the optimal binding site from a pool of duplex DNA with random sequence.

The polypeptide of the present invention has many uses. It can be used in research or in treatment to alter the expression of certain proteins by binding to the transcription control element of the protein. For example, when directed to a suitable target site in a human or animal, the polypeptide can be used to suppress the expression of a protein involved in a disease process, e.g. a viral protein or an oncogene. A zinc finger protein with high affinity and new DNA binding specificity can be used to enhance gene expression. In research, the polypeptide, when combined with a suitable label (e.g. a radioactive or fluorescent label) can be used to identify a unique polynucleotide sequence within a larger polypeptide by binding to the unique sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is the construction of the Zif phage vector fd-tet.Zif. The phage vector fUSE3 was converted into

2

fd-tet.Zif in two steps: (i) a polylinker was inserted into the Xho I site of fUSE3 and (ii) a PCR-amplified fragment of Zif268 complementary DNA (the upper strand is SEQ. ID NO. 1 and the lower strand is SEQ. ID NO. 2) encompassing bases 1287 through 1585 was cut with Apa I and Xba I and then ligated into the Apa I and Spe I sites of the polylinker.

FIG. 2 is a sketch of the Zif phage. The Zif268 zinc finger peptide, which contains three fingers (denoted by the numbered circles), is fused to the NH₂-terminal end of the phage coat protein pIII. Three to five copies of this fusion protein should be present at one end of the virion.

FIG. 3 shows the three biotinylated DNA sites used for affinity selections (SEQ ID NO. 3) and the complementary strands (SEQ. ID NO. 4). The sequences of the underscored region were GACC, GCAC, or CCTG (where XXXX indicates the appropriate complementary sequence), and these duplexes are referred to as GACC, GCAC, and CCTG. Zif normally recognizes duplex DNA with the consensus sequence GCGTGGGCG (with the first finger contacting the underlined "GCG" subsite) and the box marks the corresponding regions of the duplexes.

FIG. 4 is a overview of selections. Samples from the phage library were subjected to five rounds of selection with the biotinylated DNA duplexes GACC, GCAC, or CCTG. The GACC and GCAC pools were then used in additional rounds of selection under more stringent conditions (the washes contained more salt and, for one of the GCAC selections, the binding reactions contained nonbiotinylated Zif268 binding site as a specific competitor) (Example 4). Pools were characterized at the indicated stages by sequencing of randomly chosen phage. The X indicates that there were no further selections with the CCTG pool.

DETAILED DESCRIPTION OF THE INVENTION

One embodiment of the present invention is a polypeptide that binds a targeted polynucleotide with high affinity. "High affinity" means that the polypeptide binds the targeted polynucleotide with a dissociation constant [K_d] of less than about 1.0 nanomolar, preferably less than about 0.5 nanomolar. The polypeptide comprises three or more zinc fingers. The binding of the polypeptide to the target polynucleotide is brought about through contacts between amino acid residues of the zinc fingers and the nucleic acid bases of the polynucleotide, thereby forming a stable polynucleotide/polypeptide complex. In the present invention, the polypeptide encompassing the zinc fingers 1) differs in sequence—at one or more base-contacting amino acid residues—from a known wild type zinc finger protein or has a base sequence specificity distinct from that of a known wild type zinc finger protein and 2) binds to a targeted polynucleotide with high affinity (binds DNA with high affinity, i.e., with a dissociation constant of less than 1 nM). In one embodiment, a polypeptide of the present invention differs in the identity of its base-contacting amino acid residues from a corresponding wild type zinc finger protein, such as Zif268. In another embodiment, the polypeptide has a DNA binding specificity different from that of any known zinc finger protein. The base sequence specificity of a protein can be determined by selecting the optimal binding site from a pool of duplex DNA with random sequence.

Zinc fingers are involved in eukaryotic protein-nucleic acid interactions that control gene transcription. Naturally occurring zinc fingers, referred to herein as "wild type" zinc fingers occur in regulatory fractions that include differentiation and growth signals [EGR1, also referred to as Zif268,

(Sukhatme, V. P. et al., *Cell*, 53:37 (1988); Christy, B. A., et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:7857 (1988)), EGR2 (Joseph, L. J., et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:7164 (1988); Chavrier, P. et al., *EMBO J.* 7:29 (1988))), in protooncogenes [GLI (Kinzler, K. W., et al., *Nature*, 332:371 (1988)). Wilm's tumor gene (Call, K. M. et al., *Cell*, 60:509 (1990))), in general transcription factors [SP1 (Gidoni, D., et al., *Nature*, 312:409 (1984))], in *Drosophila* segmentation genes [Hunchback (Tautz, D., et al., *Nature*, 327:383 (1987)). Kruppel (Rosenberg, U. B., et al., *Nature*, 319:336 (1986))], and in regulatory genes of lower eukaryotic organisms [ADR1 (Bermis, L. T., and Denis, C. L., *Mol. Cell. Biol.*, 8:2125 (1988)). BrLA (Adams, T. H., et al., *Cell*, 54:353 (1988))].

As used herein, a zinc finger is a polypeptide structural motif folded around a bound zinc cation. The polypeptide of a zinc finger has a sequence of the form X₃-Cys-X₂₋₄-Cys-X₁₂-His-X₃₋₅-His-X₄, wherein X is any amino acid (e.g. X₂₋₄ indicates an oligopeptide 2-4 amino acids in length). There is generally a wide range of sequence variation in the 28-31 amino acids of the known zinc finger polypeptide. Only the two consensus histidine residues and two consensus cysteine residues bound to the central zinc atom are invariant. Of the remaining residues, three to five are highly conserved, while there is significant variation among the other residues. Despite the wide range of sequence variation in the polypeptide, zinc fingers of this type have a similar three dimensional structure. However, there is a wide range of binding specificities among the different zinc fingers, i.e. different zinc fingers bind double stranded polynucleotides having a wide range of nucleotides sequences.

The binding specificity and affinity of a zinc finger is largely determined to a large degree by the amino acid residues which contact the nucleic acids of the polynucleotide. Based on the three published reports of the x-ray crystal structures of zinc fingers complexed to DNA (Pavletich and Pabo, *Science* 252:809 (1991), Pavletich and Pabo, *Science* 261:1701 (1993) and Fariall et al., *Nature* 366:483 (1993)), there are four nucleic acid-contacting residues in zinc fingers that are primarily responsible for determining specificity and affinity. These four amino acid residues occur in the same position relative to the first consensus histidine and the second consensus cysteine. Specifically, these four amino acid residues define which three to four base pair [or subsites] the zinc finger prefers to bind (i.e. the specificity of the zinc finger) and with how great an affinity. The first of the three critical amino acid residues is seven residues to the N-terminal side of the first consensus histidine and six residues to the C-terminal side of the second consensus cysteine. This is hereinafter referred to as the "-1 position". The other three amino acids are two, three and six residues removed from the C-terminus of the residue at position -1, and are referred to as the "2 position", "3 position" and "6 position", respectively. The amino acid residues one and five residues removed from the C-terminus of the amino acid at the -1 position are also important to zinc finger specificity and binding strength. Positions one and five residues removed from the C-terminus of the amino acid at -1 are referred to as the "1 position" and "5 position", respectively. These amino acid residues at these six positions are referred to as the base-contacting amino acids. These six positions are referred to as "base contacting positions." It is to be understood that in a given zinc finger protein, not all of the amino acids at these six positions contact the double stranded DNA.

As used herein, a "targeted polynucleotide" refers to a portion of double-stranded polynucleotide acid (e.g. RNA,

DNA, PNA or combinations thereof) to which it is advantageous to bind a protein. In addition, a "targeted polynucleotide" binds wild type zinc protein with low affinity, i.e., with a binding constant greater than about 1.0 nanomolar under conditions described in Example 6. In one aspect, a "targeted polynucleotide" will be all or part of a transcriptional control element for a gene for which a desirable result can be attained by altering the degree of its expression. A transcriptional control element includes positive and negative control elements such as a promoter, an enhancer, other response elements (e.g. steroid response element, heat shock response element or metal response element) a repressor binding sites, operator and/or silencers. The transcriptional control element can be viral (e.g. the HIV promoter or the adenovirus promoter), eukaryotic (e.g., the p53 oncogene promoter) or prokaryotic. A "targeted polynucleotide" can also refer to a downstream sequence which could bind a protein and thereby prevent transcription.

For example, it is desirable to decrease the expression of oncogenes (such as mutant formed of p53) in cancer patients or certain viral genes (e.g. certain HIV or adenovirus genes) in virus-infected individuals. A protein containing zinc fingers which binds to these transcription control elements on other sites in the promoter region will cause a decrease in the expression of these genes by blocking the binding of transcription factors that normally stimulate gene expression. Consequently, the polypeptides of the present invention can be used in gene therapy. In other instances, it may be desirable to increase expression of a particular protein such as a lipoprotein receptor involved in cholesterol metabolism. A zinc protein which binds to the promoter site of the protein and which contains a transcription activator can cause such an increase in expression. In other instances, it may be desirable to attach another protein, for example a nuclease, to a zinc finger protein which binds to a double stranded polynucleotide having a particular sequence. In this case, the DNA in the vicinity of the site to which the zinc finger attaches would be degraded by the nuclease. Such a protein complex would be useful for destroying certain undesirable cells, for example cancer cells. Fragments produced by such nuclease/zinc finger complex would be useful for gene mapping and for cloning.

In another aspect, a "targeted polynucleotide" is a short portion of duplex nucleic acid (e.g. RNA, DNA, PNA or any hybrids thereof) having from about 8 to about 40 base pairs) having a defined sequence for which there is some desirable purpose in determining its presence or absence within a larger polynucleotide. For example, it may be desirable to determine whether a particular promoter or control region is found within the genome of a particular organism. A labeled protein (e.g. bound with a radioactive or fluorescent label) containing zinc fingers which binds to a polynucleotide having this particular sequence can be used to determine whether the genetic material of the organism contains this particular sequence.

When a multifinger protein binds to a polynucleotide duplex (e.g. DNA, RNA, PNA or any hybrids thereof) its fingers typically line up along the polynucleotide duplex with a periodicity of about one finger/3 bases of polynucleotide. The binding sites of individual zinc fingers (or subsites), however, typically span three to four bases, and so subsites of adjacent fingers usually overlap by one base. For example, binding of three-finger protein 123 to the 10 base pair site ABCDEFGHIJ (where these letters indicate one of the duplex DNA) could be

65 3 2 1

represented as ABCDEFGHIJ with the subsite of finger 1 being GHJ, finger 2 being DEFG and finger 3 being

ABCD. To design a three-finger protein to bind to the targeted 10 base site ABCDEFXXXX (wherein each "X" represents a base that would be specified in a particular application), zinc fingers 2 and 3 would have the same polypeptide sequence as found in wild type zinc fingers which bind DEFG and ABCD, respectively. Finger 1 would have a mutated polypeptide sequence. Preferably, finger 1 would have mutations at one or more of the base-contacting positions, i.e. finger 1 would have the same polypeptide sequence as a wild type zinc finger except that at least one of the four amino residues at the primary positions would differ. Similarly, to design a three-finger zinc protein that would bind to a 10 base sequence ABCXXXXHJJ (wherein each "X" is base that would be specified in a particular application), fingers 1 and 3 have the same sequence as the wild type zinc fingers which bind GHJJ and ABCD, respectively, while finger 2 would have residues at one or more base-coating positions which differ from those in a wild type finger. The present invention encompasses multifingered proteins in which more than one finger differs from a wild type zinc finger. It also includes multifingered protein in which the amino acid sequence in all the fingers have been changed. It is also possible to design or select a zinc finger protein to bind to a targeted polynucleotide in which more than four bases have been altered. In this case, more than one finger of the binding protein must be altered. For example, in the 10 base sequence XXXDEFGXXX, a three-finger binding protein could be designed in which fingers 1 and 3 differ from the corresponding fingers in a wild type zinc finger, while finger 2 will have the same polypeptide sequence as the corresponding finger in the wild type fingers which binds to the subsite DEFG. Binding proteins having more than three fingers can be also designed for base sequences of longer length. For example, a four finger-protein will optimally bind to a 13 base sequence, while a five-finger protein will optimally bind to a 16 base sequence. A multifinger protein can also be designed in which some of the fingers are not involved in binding to the selected DNA, and the GLI-DNA complex also shows that slight variations are possible in the spacing of the fingers.

Another embodiment of the present invention is a deoxyribonucleotide which encodes a polypeptide that binds a targeted polynucleotide with high affinity (as described above). The polypeptide comprises three or more zinc fingers. The binding of the polypeptide to the targeted polynucleotide duplex is brought about through contacts between amino acid residues of the zinc fingers and the nucleic acid bases of the targeted polynucleotide, thereby forming the duplex polynucleotide/polypeptide complex. At least one base-contacting amino acid residue in at least one zinc finger differs from the amino acid in the corresponding position in the wild type zinc finger, i.e. the polypeptide of the present invention differs from wild type zinc finger proteins in the amino residues present in at least one amino acid residue which contacts the target polynucleotide.

In a preferred embodiment, one or more amino acids in one or more zinc fingers of the polypeptide are mutated in the base-contacting positions. Examples of suitable targeted a promoter region polynucleotides include a transcription control element or promoter region, as described above. Suitable wild type zinc fingers include EGR1, EGR2, GLI, Wilson's tumor gene, Sp1, Hunchback, Kruppel, ADR1 and BrLA.

A phage display system was developed and used to select zinc finger proteins with altered DNA-binding specificities.

The three zinc fingers of the Zif268 protein were expressed on the surface of filamentous phage, and a library of variants was prepared by randomizing critical amino acids in the first zinc finger. Affinity selections, using DNA sites with base changes in the region recognized by the first finger, yielded Zif268 variants that bound tightly and specifically to the new sites. Three Zif268 zinc fingers (Christy, B. A., et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:7857 (1988)) were expressed on the surface of filamentous phage (FIGS. 1 and 2). The resulting construct-fd-tet.Zif-produced useful titers of "Zif phage" (Example 2), and these phage bound specifically to the nine-base pair site (GCGTGGGCG—the first finger contacts the underlined "GCG" subsite) recognized by Zif268. A library of Zif variants were created by randomizing the four positions of the first finger that appear most important for making base contacts (Pavletich, N. P. and Pabo, C. O., *Science*, 261:1701 (1993); Pavletich, N. P. and Pabo, C. O., *Science*, 252:809 (1991); Fairall, L., et al., *Nature*, 366:483 (1993)). These randomized positions include the residue immediately preceding the a helix and include the second, third, and sixth residues of the helix (Example 3). [Note: These correspond to positions -1, 2, 3 and 6 that were described previously.]

Affinity selection methods were then used to search the library for phage that would recognize altered binding sites. In each round of affinity selection, phage were equilibrated with biotinylated target DNA and then applied to streptavidin-coated microtiter wells. After washing, the retained phage were eluted in high salt buffer, amplified in *Escherichia coli*, and purified to prepare for the next cycle. The target DNA duplexes for these selections contained modified Zif268 binding sites with changes in the region recognized by finger one (Pavletich, N. P. and Pabo, C. O., *Science*, 252:809 (1991)), and each duplex is referred to by the sequence of this region (FIG. 3). Initially, five rounds of selection were performed with each of the target sites (FIG. 4) (Example 4). During these initial selection series, retention efficiencies in the GACC- and GCAC-selected phage pools increased about 100 times, whereas retention efficiencies for the CCTG pool remained low (Example 5). These enriched GACC and GCAC were then used pools as a starting point for additional, more stringent selection cycles (FIG. 4) (14). The CCTG pool was not studied further.

Phage pools from critical stages of the GACC and GCAC selections were characterized by sequencing (FIG. 4), and amino acid preferences were apparent in each pool. For the GACC pool, sequencing after the initial selection series showed that all of the phage (12/12) could be characterized by the consensus sequence (S/D/T)_NR (Table 1). Three additional rounds of selection using high salt washes did not substantially change this consensus (Table 2). For the GCAC selections, sequencing revealed notable changes in the later pools. After the initial selection series, many of the phage belonged to a group characterized by the consensus sequence R_DR (18/22), but there also was a group characterized by the sequence _G(S/T)R (4/22) (Table 1). After additional rounds of selection with high salt washes, a single sequence RADR—from the first group predominated (Table 2). However, when the additional rounds of selection used both high salt washes and competitor Zif268 site in the binding reactions, a single sequence -QGSR- from the second group predominated.

TABLE 1

GACC -1 2 3 6	GCAC -1 2 3 6
SQNR (4,2)	RSDR (4,2)
DANR (2,1)	RPDR (3,2)
DRNR	RGDR (3,1)
DSNR	HSDR (2,2)
SSNR	RVDR (2,2)
STNR	AADR
TANR	KSDR
TPNR	RADR
	RAER
	<u>R</u> DR
	NGSR (2,2)
	SGST
	TGTR
S/D/R NR	<u>G</u> S/T R

Amino Acid sequences of phage from the GACC and GCAC pools after the initial selection series (FIG. 4). The four randomized positions in the α -helical region of finger one are denoted as -1, 2, 3, and 6. Consensus sequences are indicated in bold. An underscore () indicates that there is no clear preference at the corresponding position. The numbers in parentheses indicate the total number of times this amino acid sequence was recovered and the number of distinct DNA sequences that encoded this amino acid sequence. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; G, Gly; H, His; K, Lys; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; and V, Val.

TABLE 2

GACC -1 2 3 6	GCAC -1 2 3 6	GCAC (+ competitor) -1 2 3 6
DSNR (8,4)	RADR (7,4)	QGSR (16,3)
SSNR (4,3)		
DRNR (2,1)		
NSNR		
D/S SNR	RADR	QGSR

Amino acid sequences in the final phage pools (after the later selection series shown in FIG. 4). The designation "+ competitor" indicates that specific competitor DNA (nonbiotinylated wild-type Zif268 site) was added to the binding mixes during the later selection series (Example 4). Symbols are as described in Table 1.

Three Zif268 variants were studied in more detail by recloning and overexpressing them in *E. coli*, then purifying the resultant peptides and measuring DNA-binding affinities. Also studied were the predominant Zif268 variants obtained in each of the three later selection series—DSNR, RADR, and QGSR (Table 2)—and also included the wild-type peptide RDER as a control. Peptide affinities for each of three binding sites [GACC, GCAC, and GCGC (wild type)] were determined by quantitatively gel-shift analysis (Table 3) (Example 6). Each of the variant peptides binds with high affinity to the site used for its selection (Table 3, boxed entries). Moreover, the DSNR and QGSR peptides exhibit new specificities in that they bind to these new sites substantially better than they bind to GCGC. The RADR peptide (unlike the QGSR peptide) does not discriminate well between GCAC and GCGC. The only difference in the selection conditions for these two variants was the use of competitor Zif268 site in the selections that yielded QGSR.

TABLE 3

Finger one Sequence	Apparent K _d (nM)			
	-1 2 3 6	GACC	GCAC	GCGC (wt)
DSNR		0.019	2.5	1.8
RADR		9.3	0.068	0.035
QGSR		1.8	0.055	0.54
RDER (wt)		33.0	5.6	2.7

Apparent K_d's for the binding of zinc finger peptides to DNA fragments containing the GACC, GCAC, and GCGC (wild type) forms of the Zif268 binding site. Each peptide is specified by the amino acid residues at the four positions of finger one that were randomized in the library (-1, 2, 3, and 6). 15 RDER is wild type. The three DNA duplexes share the sequence (SEQ ID NO 5) and are specified by the bases at the positions of four X's. [The bracketed region marks the position of the Zif268 binding site GCGTGGGG (SEQ ID NO:9) (Pavletich, N. P. and Pabo, C. O., *Science*, 252:809 (1991)).

The experiments reported herein demonstrate that the 20 phage display system can be used to select zinc fingers with novel DNA-binding specificities. Based on the present disclosure, the skilled artisan will be able to select a wide variety four base pair sequences and engineer zinc fingers which bind specifically to desired four base pair sequence.

Based on the present disclosure, it is possible to construct 25 a multifingered protein in which more than one finger has each been selected to bind to a specific subsite. As a result, a multifingered protein can be engineered such that each finger binds to an adjacent and overlapping subsite. For example, a three-fingered protein can be engineered to bind 30 to the underlined target in the HIV promoter site: ATGCT-GCATATAAGCAGCTGCTT (SEQ. ID NO 6). The underlined target contains three overlapping subsites (CTGC, CATA and ATAA). A zinc finger phage library is prepared as in Examples 1-3. DNA for selection by zinc finger phage is 35 then prepared by substituting XXXX in FIG. 3 with one of the overlapping subsites, for example CTGC. Zinc finger phage is then selected for binding to the selection DNA, as in Example 4, to obtain a pool of phages expressing zinc fingers specific for CTGC. The DNA coding for the zinc fingers which bind CTGC are then amplified by polymerase 40 chain reaction to give a first pool of finger genes. This process is repeated for the other two subsites to give a second and third pool of finger genes. The three pools are then used to construct a new phage vector in which each of three zinc fingers have been selected for binding to the 45 adjacent and overlapping subsite (i.e. the first finger will contact ATAA, the second finger will contact CATA, and the third finger will contact CTGC). Phages expressing this new library will then be selected for binding to the HIV promoter target sequence. A three-fingered protein will be obtained 50 which binds to the HIV target sequence.

The invention is illustrated by the following examples, which are not to be construed as limiting in any way.

EXAMPLE 1

Construction of the Zif Phage Vector fd-tet-Zif

The phage vector fUSE3 (19) was converted into fd-tet-Zif in two steps: (i) A polylinker was inserted into the Xho I site of fUSE3 and (ii) a PCR-amplified fragment of Zif268 complementary DNA encompassing bases 1287 through 1585 (Christy, B. A., et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:7857 (1988)) was cut with Apa I and Xba I and then ligated into the Apa I and Spe I sites of the polylinker.

EXAMPLE 2

Production of Zif Phage

MC1061 cells (Smith, G. P. and Scott, J. K., *Methods Enzymol.*, 217:228 (1993)) that contained fd-tet.Zif were

grown to saturation (room temperature, about 40 hours, no agitation) under anaerobic conditions in 0.05 liter of Zif phage broth [(tryptone (32 g liter⁻¹), yeast extract (20 g liter⁻¹), glucose (8 g liter⁻¹), casamino acids (1 g liter⁻¹) tetracycline (0.02 g liter⁻¹), thiamine (0.5 mg liter⁻¹), 100 μ M ZnCl₂, 50 μ M dithiothreitol, 86 mM NaCl and 40 mM Hepes (pH 7.8)]. To increase titer, cultures were grown in dialysis tubes (50-kD cutoff) suspended in 1 liter of broth. Phage were titered as tetracycline-transducing units (TTU) essentially as described (Smith, G. P. and Scott, J. K., *Methods Enzymol.*, 217:228 (1993)) except that starved K91 cells had been stored at -80° C. in a buffer containing 67 mM NaCl, 42 mM NH₄H₂PO₄ and 14% glycerol. Titers of Zif phage cultures were 0.5×10⁹ to 14×10⁹ TTU/ml.

EXAMPLE 3

Construction of the Phage Library

Two oligonucleotides were synthesized: 5'-GGAATCGATTCCATGGGCCCATGAACGGCC GTACGCCCTGCCCTGTCGAGTCCTGCGATCGT CGATTTTCG (SEQ ID NO. 7) and 5'-CCATCTCGATC GCATGCATA TTCCGACACTGGAAAGGGCTCTGGCCT GTGTGGATCCGGATATGSNNGGTGAGSNNSNNA GASNNCGAAAATCGACG (SEQ ID NO. 8) (N=A, T, G and C; S=G and C), with complementary 12-base 3' ends. These were annealed and then extended with sequenase 2.0 (United States Biochemical). The resulting duplex was digested with Apa I and Sph I (sites are underlined) and ligated with the large Apa I-Sph I fragment of fd-tet.Zif. Ligation products were electroporated into MC1061 cells (Smith, G. P. and Scott, J. K., *Methods Enzymol.*, 217:228 (1993)), and this yielded about 2.8×10⁷ independent transformants. This library was grown essentially as described in Example 2. Phage were purified by ultracentrifugation (171, 000 g, 4° C., 6 hours), and phage pellets were resuspended in about 1/100 volume of binding buffer [50 mM NaCl, 5 mM MgCl₂, 10 μ M ZnCl₂, 5% glycerol, bovine serum albumin (BSA; 0.1 mg/ml), and 15 mM Hepes (pH 7.8)]. This final phage library preparation (about 4.7×10¹¹ TTU) was stored anaerobically (<1 ppm O₂) on ice. Because of concerns about oxidation, all phage manipulations were done so as to minimize exposure to O₂. To estimate library complexity, 20 unselected were sequenced clones. (Single-stranded templates were sequenced with sequenase 2.0 and protocols from United States Biochemical). Three corresponded to the parent construct (fd-tet.Zif) and appear to have resulted from the reinsertion of the fragment excised during library construction. Seventeen phage contained the correct library insert, but there was a significant cytosine bias at the randomized codons. Base ratios were C:A:T:G= 48:19:19:15 at the first two codon positions and C:G=74:6 at position 3.

EXAMPLE 4

Selection Protocol for Phage which Recognize Altered Binding Sites

The selection protocol is based on the "biopanning" procedure (Smith, G. P. and Scott, J. K., *Methods Enzymol.*, 217:228 (1993)). The first round in each initial selection series (the leftmost arrow in each of the three pathways in FIG. 4) was done as follows: Binding reactions (121 μ l) were made that contained about 3.5×10¹⁰ TTU of library phage, 39 nM of biotinylated target DNA [GACC, GCAC, or CCTG (FIG. 3)], and sheared calf thymus DNA (0.059 mg/ml) in 0.9× binding buffer (see Example 3). Each sample

was preincubated for 50 minutes, diluted into 3.6 volumes of 0.05 M NaCl wash buffer [0.05 M NaCl with 5 mM MgCl₂, 10 μ M ZnCl₂, 5% glycerol, 0.5% w/v Triton X-100, and 15 mM Hepes (pH 7.8)], and applied to streptavidin-coated wells (six wells, 30 μ l per well) of a Pro-Bind plate (Becton Dickinson). After 50 minutes the samples were removed from the wells and then (i) over a period of 35 minutes, the wells were rinsed 10 times with 0.25 ml of 0.45M NaCl wash buffer (identical to 0.05M NaCl wash buffer except for the higher NaCl concentration) and (ii) 40 μ l of elution buffer [binding buffer (see Example 3) with 4M NaCl] was added to each well. After eluting for 2 hours, each set of six eluates was pooled, titered, and used to infect K91 cells (see Example 2). Transduced cells were incubated for 1 hour at 37° C. in 5 ml of LB broth containing tetracycline (0.2 μ g/ml), centrifuged (15 minutes, 1600 g), and resuspended in 50 ml of degassed Zif phage broth. Each culture then grown anaerobically in a 50-ml centrifuged tube and purified essentially as described in Example 3. Other rounds of selection in the initial series were similar except that, starting at round 3, sonicated salmon sperm DNA was substituted for sheared calf thymus DNA in the binding reactions. Selections in the later series were similar except that 0.75M NaCl washes were used and the binding reactions in one of the GCAC selection series included a nonbiotinylated Zif268 binding site (0.36 μ M) as a specific competitor. All phage manipulations, except for elution and infection of K91 cells, were carried out in an anaerobic chamber with less than 1 ppm of O₂.

In the first round of selection, <0.009% of library phage applied to the streptavidin-coated wells was recovered in the eluates. By the fifth round this retention efficiency had risen to 0.6 to 0.8% for the GACC and GCAC phage pools, but was less than 0.001% for the CCTG pool. For comparison, control experiments using Zif phage and a biotinylated wild-type Zif site typically yielded retention efficiencies of 0.5 to 1.0%.

EXAMPLE 5

Purification of Zinc Finger Variants

The zinc finger regions from the phage variants [corresponding to residues 333 to 421 of Zif268 (Christy, B. A., et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:7857 (1988))] were subcloned into the T7 expression vectors pET-3d or pET-21d (Novagen). These expression constructs were transformed into BL21 cells containing the pLysS plasmid and then induced as recommended (Novagen). Additionally, the corresponding wild-type peptide (RDER) was expressed as described (Pavletich, N. P. and Pabo, C. O., *Science*, 252:809 (1991)). Zinc finger peptides were purified by reversed-phase batch extraction and reversed phase high-performance liquid chromatography (Pavletich, N. P. and Pabo, C. O., *Science*, 252:809 (1991)). The final peptide preparations were reconstituted in water in an anaerobic chamber and adjusted to 2.75 mM ZnSO₄ and 50 mM bis-tris propane (pH 6.8). Peptide samples were stored at -80° C. To estimate purity, peptides were subjected to SDS-polyacrylamide gel electrophoresis and silver staining. No impurities staining as intensely as 2% of the purified peptide were observed in any preparation (12).

EXAMPLE 6

Derivation of Binding Constant

To derive apparent dissociation constants (K_d 's) (i) quantitative gel-shift analysis was used to determine the fraction

11

of DNA fragment bound at a series of peptide concentrations, (ii) the K_d was estimated at each point in the transition region of the resulting "binding curve," and (iii) these K_d 's were averaged. Those points were used for which $0.1 \leq \text{fraction DNA bound} \leq 0.9$ (six or seven points). Standard deviations were always $< K_d(\text{average})/4$. Binding reactions contained radioactive DNA fragment (about 2.5 pM or about 25 pM), peptide (from a twofold dilution series), and poly (dl-dC)-poly (dl-dC) (14.7 $\mu\text{g}/\text{ml}$; Pharmacia) in degassed gel-shift buffer [50 mM NaCl, 5 mM MgCl₂, 10 μM ZnSO₄, 5% glycerol, BSA (0.1 mg/ml), 0.1% NP-40, and 15 mM Hepes (pH 7.8)]. Binding reactions were equilibrated at room temperature for either 30 minutes (for RDER) or 4 hours (for the variant peptides) and electrophoresed on 10% poly-acrylamide gels in 0.03M tris-Hepes (pH 7.8). (Control experiments showed that the variant peptides required longer equilibration times). Dried gels

12

were quantitated with the use of a Phosphorimager system (Molecular Dynamics). A freshly thawed sample of peptide was used for each set of gel-shift experiments, and the binding activity was determined by titrating a portion of each sample against a defined concentration of binding site (150 μM or 300 μM). Each sample was titrated twice, with two different DNA fragments (of the three in Table 3), and the calculated activities always agreed within 20%.

10

15

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described specifically herein. Such equivalents are intended to be encompassed in the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 9

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGATCAGGG CCCTTCGAAA CTAGTGCACC TGCAGAGGAT CCT

4 3

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCGAAAGATC CTCTGCAGGT GCACTAGTT CGAAGGGCCC TGA

4 3

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGAGCGTGG NNNNAGTGAT CGATC

2 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CACTNNNNCC ACGCTCAG

1 8

5,789,538

13

14

-continued

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGCAGCTGAG CGTCCNNNN GTGAGCT

27

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATGCTGCATA TAAGCAGCTG CTT

23

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAATCGATT CCATGGGGCC CCATGAACGG CCGTACGCC TGCCTGTG AGTCCTGCGA

60

TCGTCGATTT TCG

73

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCATCTCGAT CGCATGCATA TTGACACTG GAAGGGCTTC TGGCTGTGT GGATCCGGAT

60

ATGSNNNGTG AGSNNSNAG ASNNCGAAAA TCGACG

96

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCTGGGGCG

9

What is claimed is:

1. A polypeptide comprising three zinc fingers, each of said zinc fingers comprising a set of base-contacting amino acid residues, wherein:

a) at least one set of base-contacting amino acid residues in at least one zinc finger differs in sequence at one or

65

more base-contacting amino acid residues as compared with a corresponding set of base-contacting amino acid residues in any known wild type zinc finger protein; b) the polypeptide has a DNA base sequence specificity different from that of any known wild type zinc finger protein; and

15

c) the polypeptide binds a targeted polynucleotide with a dissociation constant of less than about 1.0 nanomolar.

2. The polypeptide of claim 1 wherein the three or more zinc fingers bind the targeted polynucleotide with an dissociation constant of less than about 0.5 nanomolar.

3. The polypeptide of claim 2 wherein the wild type zinc finger is selected from the group consisting of EGR1, EGR2, GLI, Wilms tumor gene, Sp1, Hunchback, Kruppel, ADR1 and BrLA.

4. The polypeptide of claim 1 wherein the targeted polynucleotide is an HIV promoter.

5. The polypeptide of claim 1 wherein the amino acids at positions 3 and 6 of at least one zinc finger are asparagine and arginine, respectively, wherein the amino acid at position -1 is selected from the group consisting of serine, 15 threonine and aspartic acid and wherein the targeted polynucleotide comprises GACC.

6. The polypeptide of claim 1 wherein the amino acids at positions -1, 3 and 6 of at least one zinc finger are arginine.

16

aspartic acid and arginine, respectively, and the targeted polynucleotide comprises GCAC.

7. The polypeptide of claim 1 wherein the amino acids at positions 2 and 6 of at least one zinc finger are glycine and arginine, respectively, and wherein the amino acid at position 3 is serine or threonine and wherein the targeted polynucleotide comprises GCAC.

8. A polypeptide comprising three zinc fingers, wherein the polypeptide

a) has essentially the same amino acid sequence as a wild type zinc finger protein, further characterized by at least one or more point mutations at one or more base-contacting amino acid residues;

b) has a base sequence specificity different from the wild type zinc finger protein; and

c) binds a targeted polynucleotide with a dissociation constant of less than about 1.0 nanomolar.

* * * * *



US006013453A

United States Patent [19]

Choo et al.

[11] Patent Number: 6,013,453
[45] Date of Patent: Jan. 11, 2000

[54] BINDING PROTEINS FOR RECOGNITION OF DNA

[75] Inventors: Yen Choo, Singapore, Singapore; Aaron Klug, Cambridge, United Kingdom; Isidro Sanchez Garcia, Salamanca, Spain

[73] Assignee: Medical Research Council, London, United Kingdom

[21] Appl. No.: 09/139,762

[22] Filed: Aug. 25, 1998

Related U.S. Application Data

[63] Continuation of application No. 08/793,408, filed as application No. PCT/GB95/01949, Aug. 17, 1995.

Foreign Application Priority Data

Aug. 20, 1994 [GB] United Kingdom 9416880
Nov. 8, 1994 [GB] United Kingdom 9422534
Jul. 19, 1995 [GB] United Kingdom 9514698

[51] Int. Cl.⁷ C12Q 1/68; C07H 21/04

[52] U.S. Cl. 435/6; 536/23.4

[58] Field of Search 435/6; 536/23.4

[56] References Cited

U.S. PATENT DOCUMENTS

5,498,530 3/1996 Schatz et al. 435/69.1

OTHER PUBLICATIONS

Rebar et al: "Zinc Finger Phage: Affinity Selection of Fingers with New DNA-Binding Specificities", *Science*, vol. 263, Feb. 4, 1994, pp. 671-673.

Jamieson et al: "In Vitro Selection of Zinc with Altered DNA-Binding Specificity", *Biochemistry*, 1994, 33, pga 5689-5695.

Thiesen et al: "Determination of DNA binding specificities of mutated zinc finger domains", *FEBS Letters*, vol. 283, No. 1, May 1991, pp. 23-26.

Jacobs: "Determination of the base recognition positions of zinc fingers from sequence analysis", *The EMBO Journal*, vol. 11, No. 12, 1992, pp. 4507-4517.

Desjarlais et al: "Toward rules relating zinc finger protein sequences and DNA binding site preferences" *Proc.Natl.Acad.Sci.USA*, vol. 89, Aug. 1992, *Biophysics*, pp. 7345-7349.

Nardelli et al: "Zinc finger-DNA recognition: analysis of base specificity by site-directed mutagenesis", *Nucleic Acids Research*, vol. 20, No. 16, pp. 4137-4144.

Desjarlais et al: "Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins", *Proc.Natl.Acad.Sci.USA*, vol. 90, Mar. 1993, *Biochemistry*, pp. 2256-2260.

Choo et al: "Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage", *Proc.Natl.Acad.Sci.USA*, vol. 91, Nov. 1994, *Biochemistry*, pp. 11163-11167.

Choo et al: "Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions", *Proc.Natl.Acad.Sci.USA*, vol. 91, Nov. 1994, *Biochemistry*, pp. 11168-11172.

Choo et al: "In Vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence", *Nature*, vol. 372, Dec. 15, 1994, pp. 642-645.

Wu et al: "Building zinc fingers by selection: Toward a therapeutic application", *Proc.Natl.Acad.Sci.USA*, vol. 92, Jan. 1995, *Biochemistry*, pp. 344-348.

Klug et al: "Zinc Fingers", *The FASEB Journal*, May 1995, vol. 9, No. 8, pp. 597-604.

Choo et al: "Designing DNA-binding proteins on the surface of filamentous phage", *Current Opinion in Biotechnology*, 1995, 6: pp. 431-436.

Orkin, Report and recommendation of the panel to assess the NIH investment in research on gene therapy, Dec. 7, 1995.

Freisen et al. Phage display of RNA binding zinc fingers from transcription factor IIIA, *J. Biol. Chem.* vol. 272(17): 10994-10997, Apr. 25, 1997.

Blaese et al., Vectors in cancer therapy: how will they deliver?, *Cancer Gene Therapy*, vol. 2(4): 291-297, Oct. 1995.

Primary Examiner—Nancy Degen

Assistant Examiner—William Sandals

Attorney, Agent, or Firm—Pillsbury Madison & Sutro LLP

[57] ABSTRACT

Disclosed are libraries of DNA sequences encoding zinc finger binding motifs for display on a particle, together with methods of designing zinc finger binding polypeptides for binding to a particular target sequence and, inter alia, use of designed zinc finger polypeptides for various in vitro or in vivo applications.

26 Claims, 14 Drawing Sheets

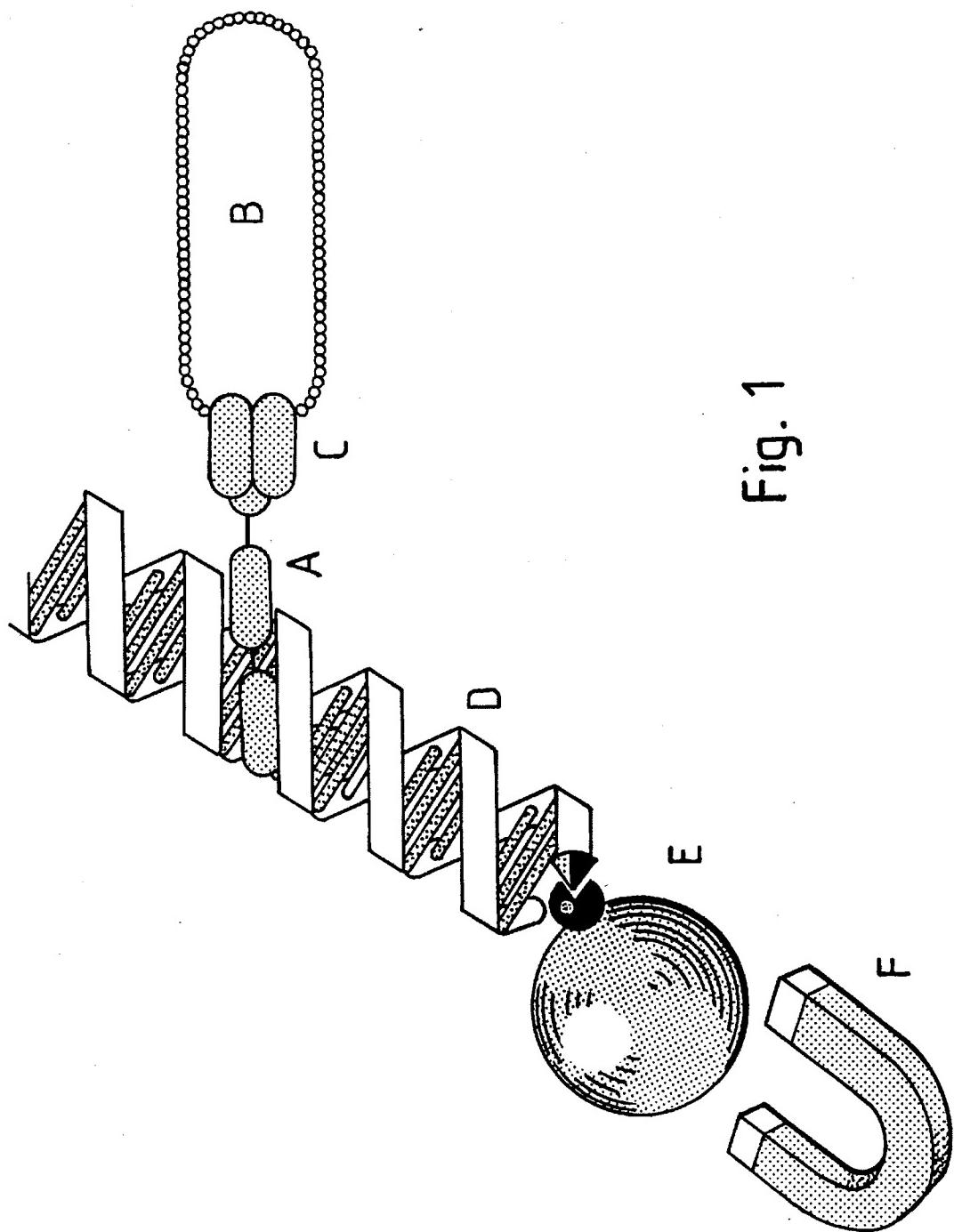


Fig. 1

- 1 1 2 3 4 5 6 7 8 9

M A E E R P Y A C P V E S C D R R F S R S D E C L T R C T R I C T
G Q K P F Q C R I - - C M R N F S X X X X C L X X H X K T H T
G E K P F A C D I - - C G R K F A R S D E R K R C T K I C L R Q K D

β

α

Fig. 2

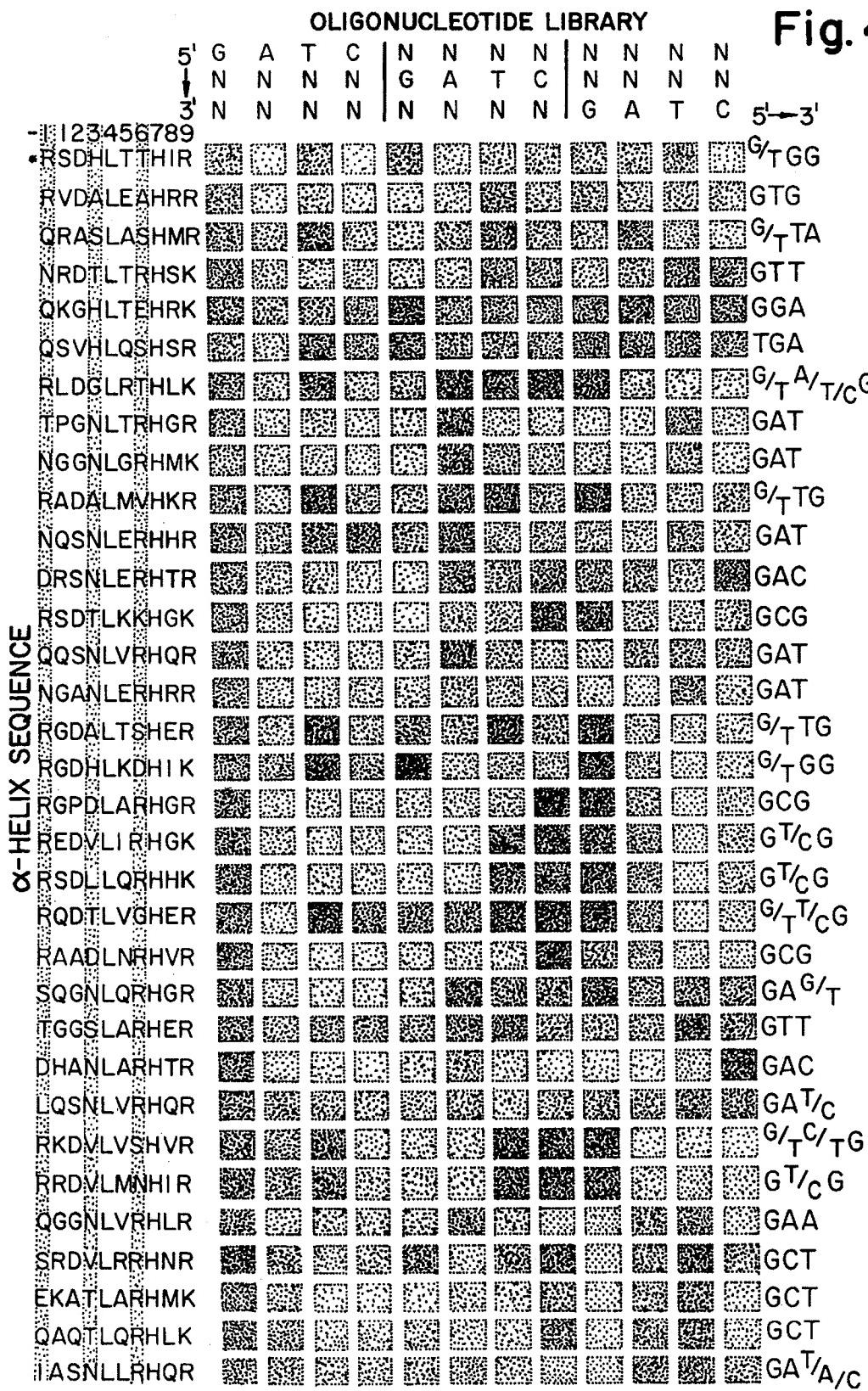
(i) TATGACTTGGATGGGAGACCGCCCTGG
 ACTGAACCTACCCTCTGGGGACCTAA -

(ii) TATATA^GGTGGGGTATATA
 ATATATGCCCATATATGCG -

(iii) TATATA^GXXXGGGTATATA
 ATATATGCCXXXGGCATATATGCG -

Fig. 3

Fig. 4



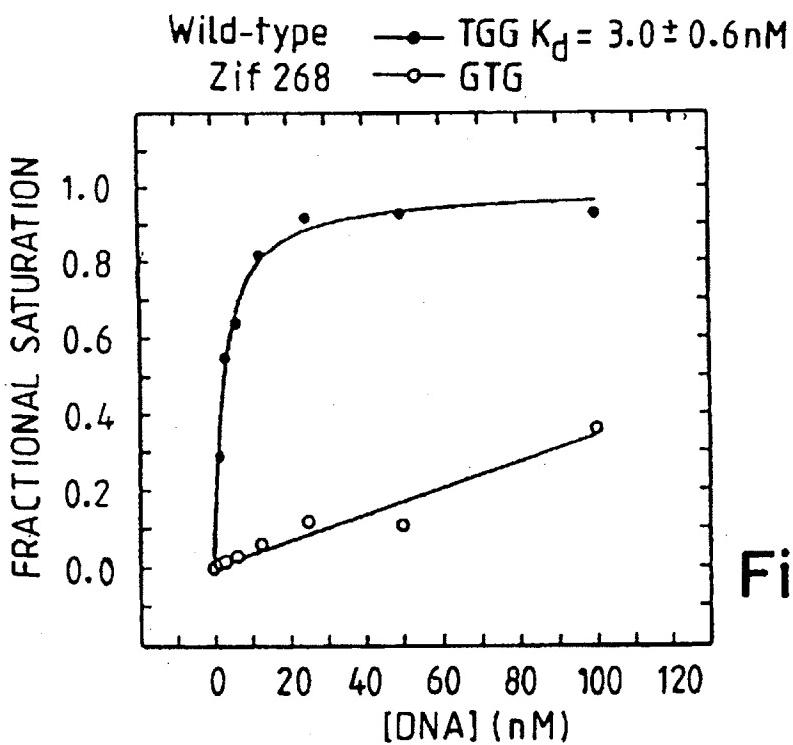


Fig. 5A

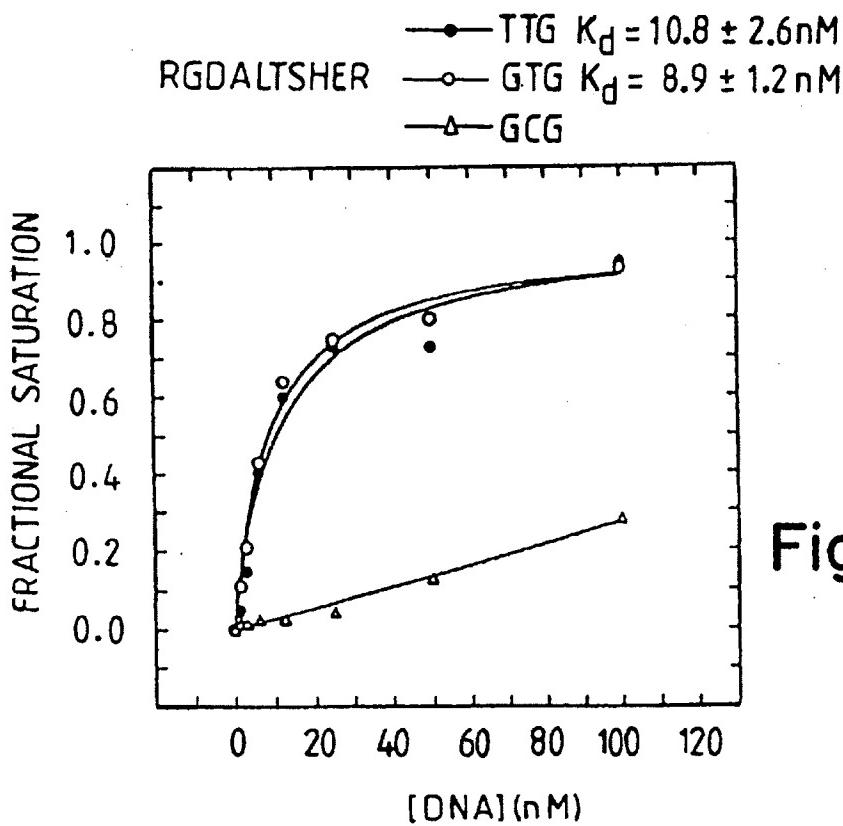


Fig. 5B

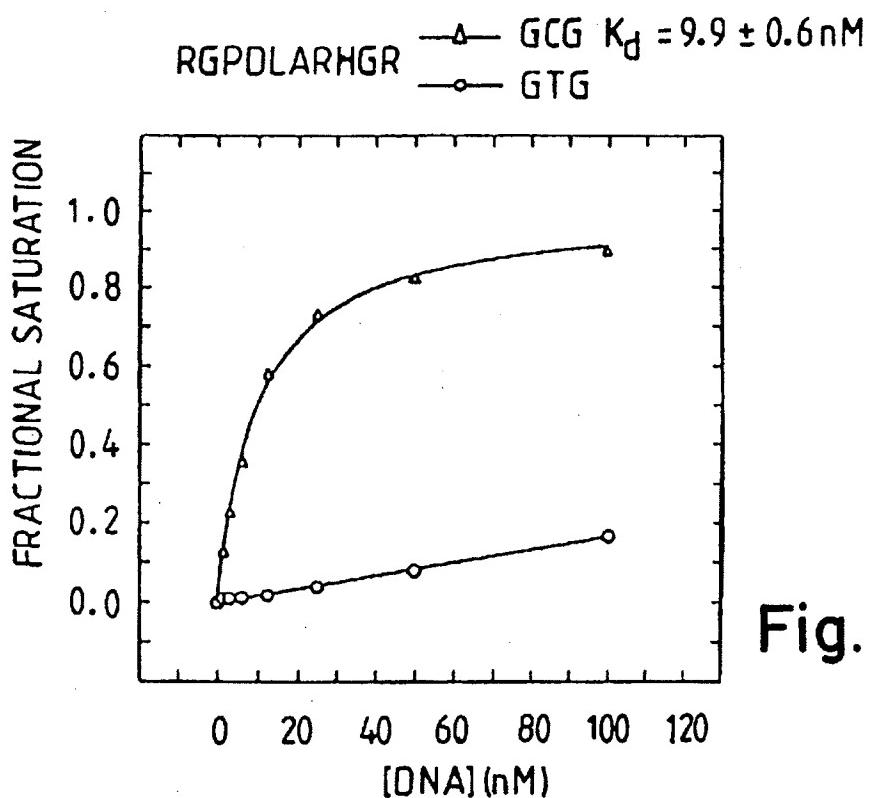


Fig. 5C

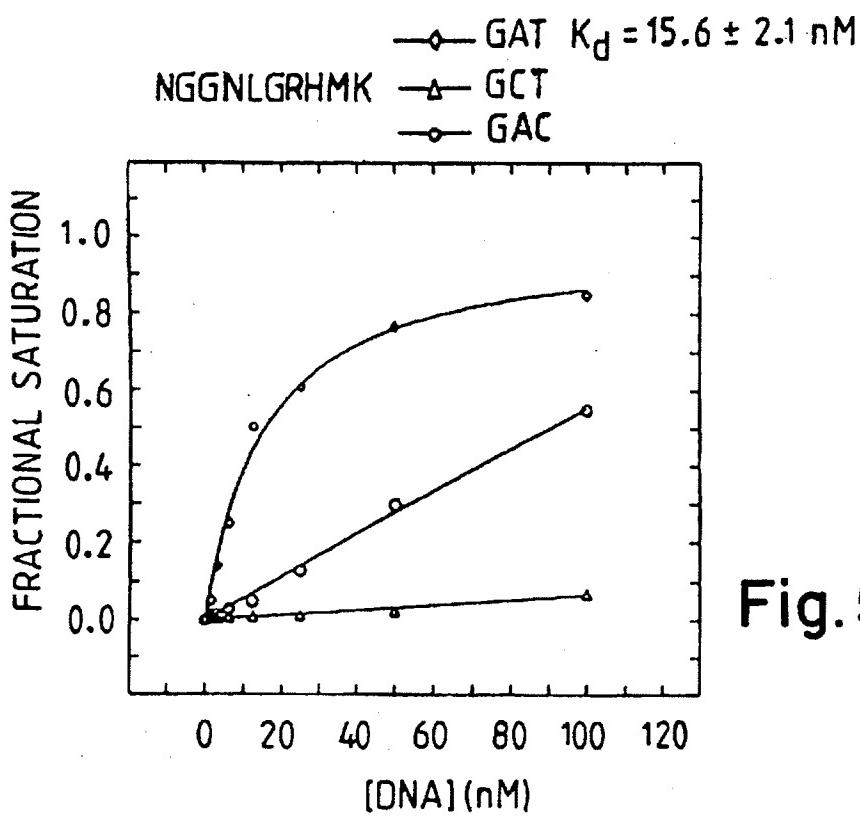


Fig. 5D

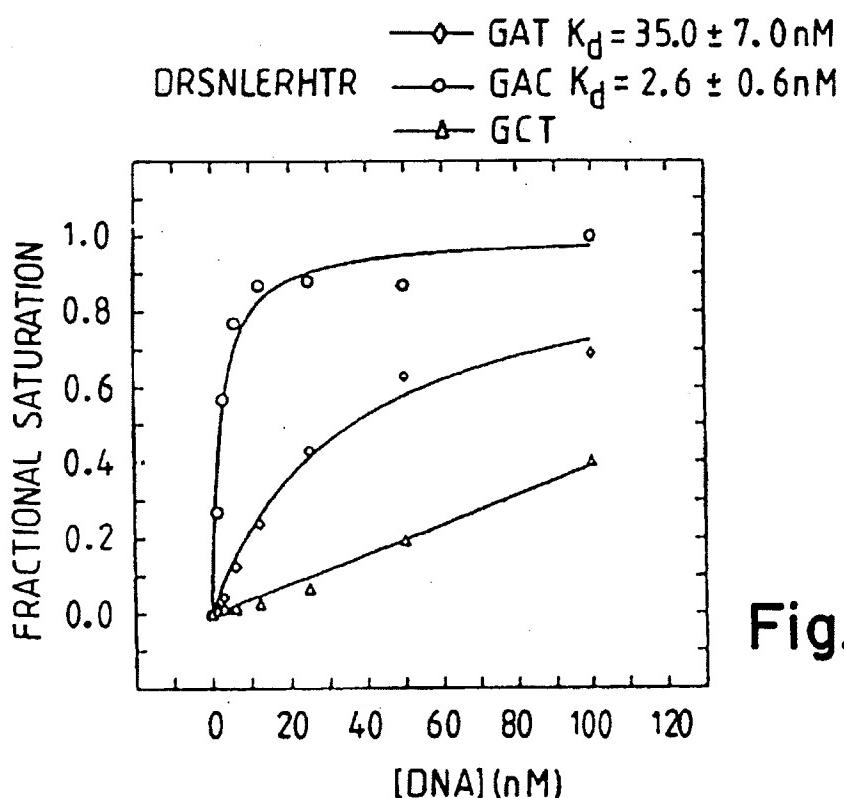


Fig. 5E

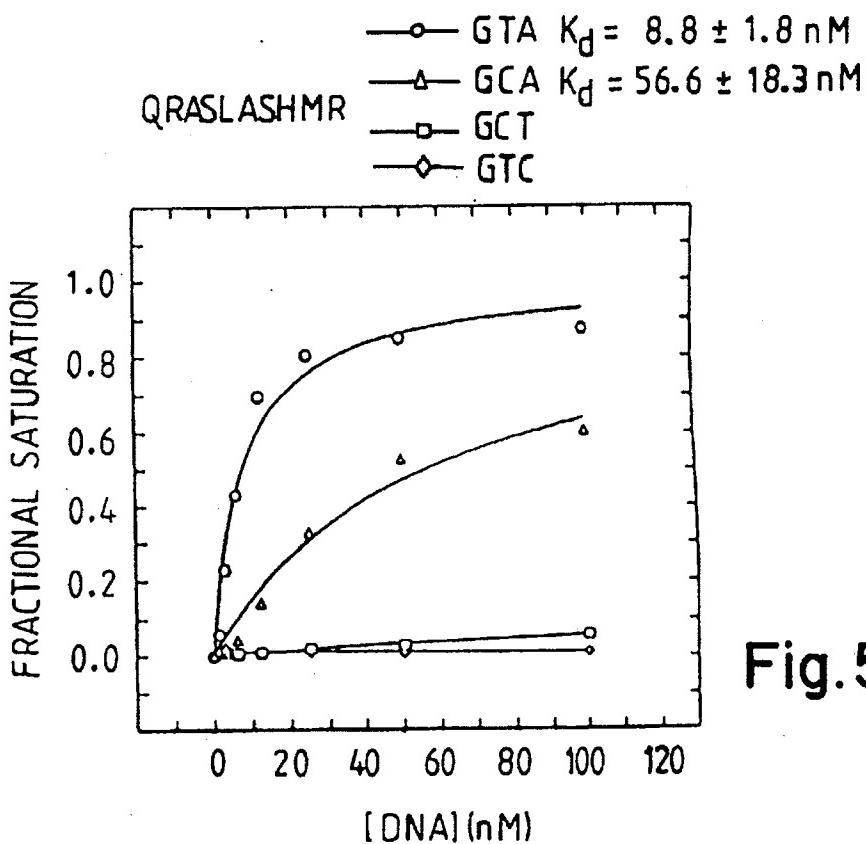


Fig. 5F

1	<i>BCR - ABL</i>	TTC	CAT	GGA	GAC	<u>GCA</u>	<u>G</u>	<u>AA</u>	<u>GCC</u>	<u>CTT</u>	<u>CAG</u>	<u>CGG</u>	<u>CCA</u>
2	<i>BCR</i>	TTC	CAT	GGA	GAC	<u>GCA</u>	<u>G</u>	<u>gt</u>	<u>gag</u>	<u>tgc</u>	<u>ctc</u>	<u>acg</u>	<u>cca</u>
3	<i>ABL</i>	ccc	ctt	tct	ttt	<u>cca</u>	<u>g</u>	<u>AA</u>	<u>GCC</u>	<u>CTT</u>	<u>CAG</u>	<u>CGG</u>	<u>CCA</u>

Fig.6

-1 1 2 3 4 5 6 7 8 9 T G E K P

1A M A E E K P F Q C R I C M R N F S D R S S L T R H T R H T G E K P

1B M A E E K P F Q C R I C M R N F S E R G T L A R H E K E T G E K P

2A F Q C R I C M R N F S Q G G N L V R H L R H T G E K P

3A F Q C R I C M R N F S Q A Q T L Q R H L K H T G E K

3B F Q C R I C M R N F S Q A A T L Q R H L K H T G E K

3C F Q C R I C M R N F S Q A Q D L Q R H L K H T G E K

β-SHEET β-SHEET α-HELIX LINKER

Fig. 7

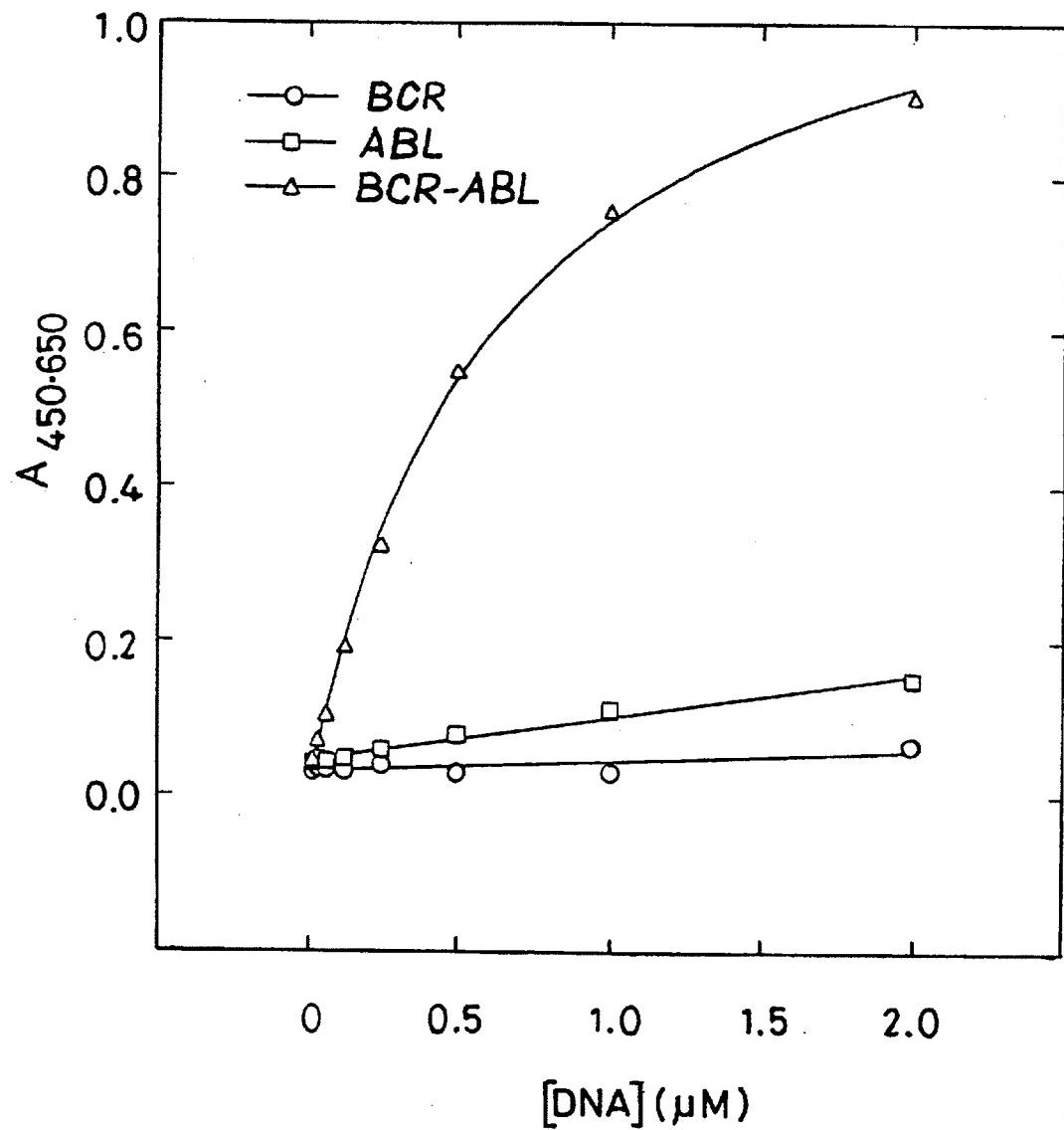


Fig. 8

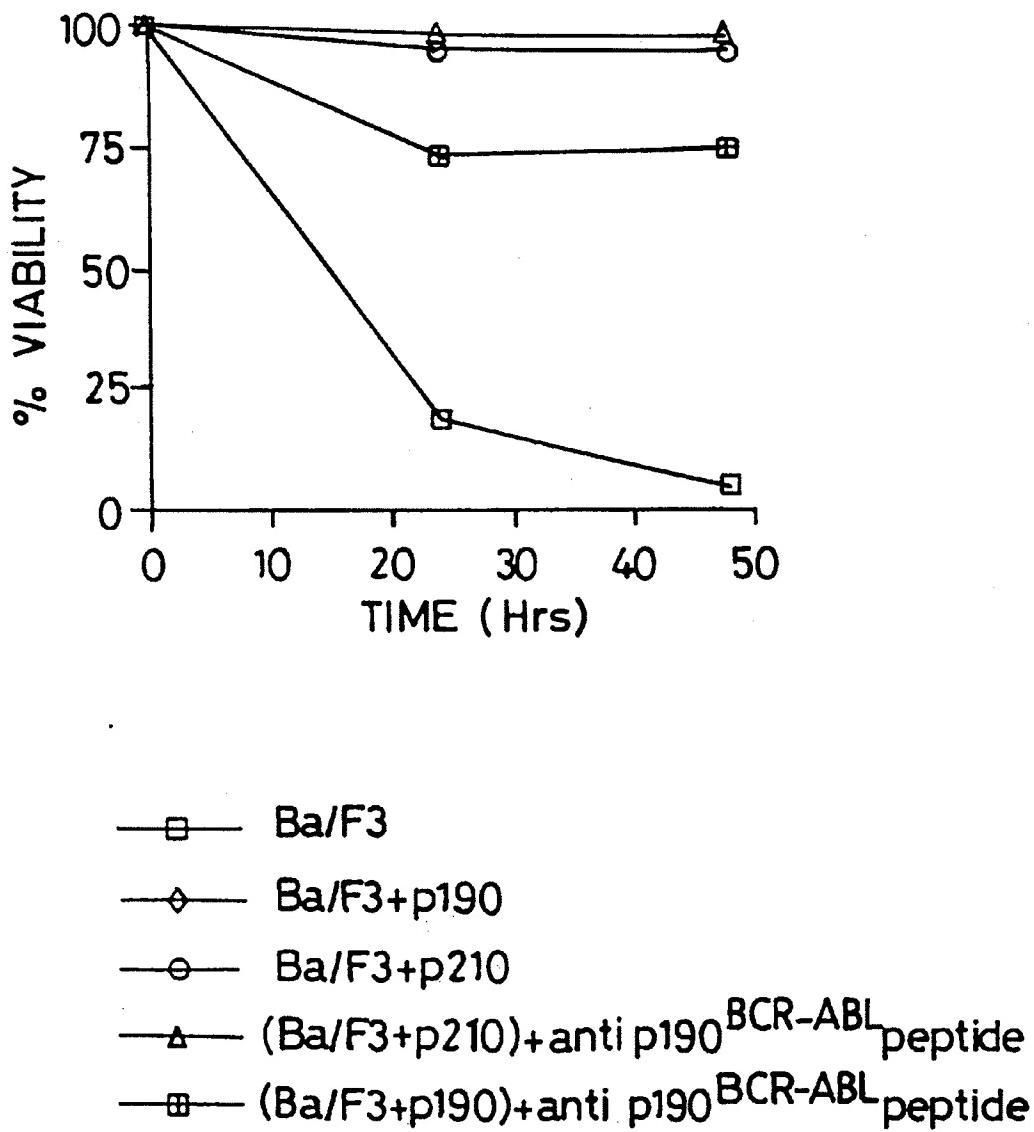


Fig. 9

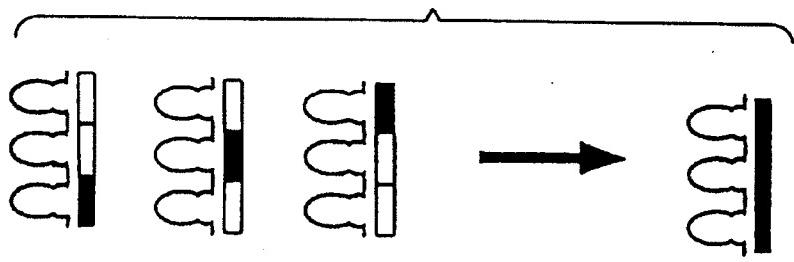


Fig. 10C

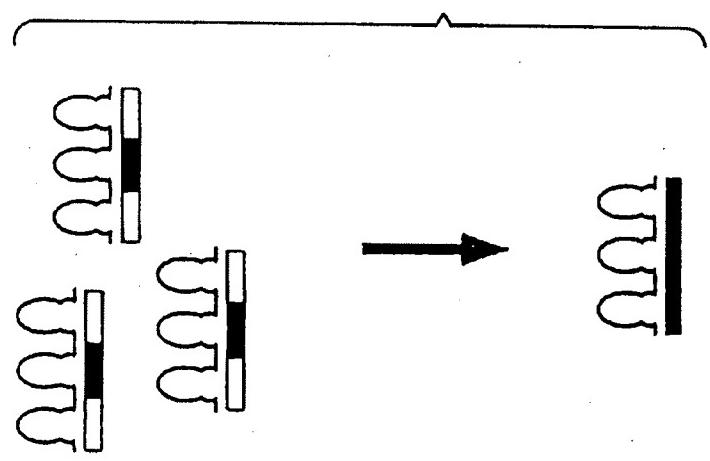


Fig. 10B



Fig. 10A

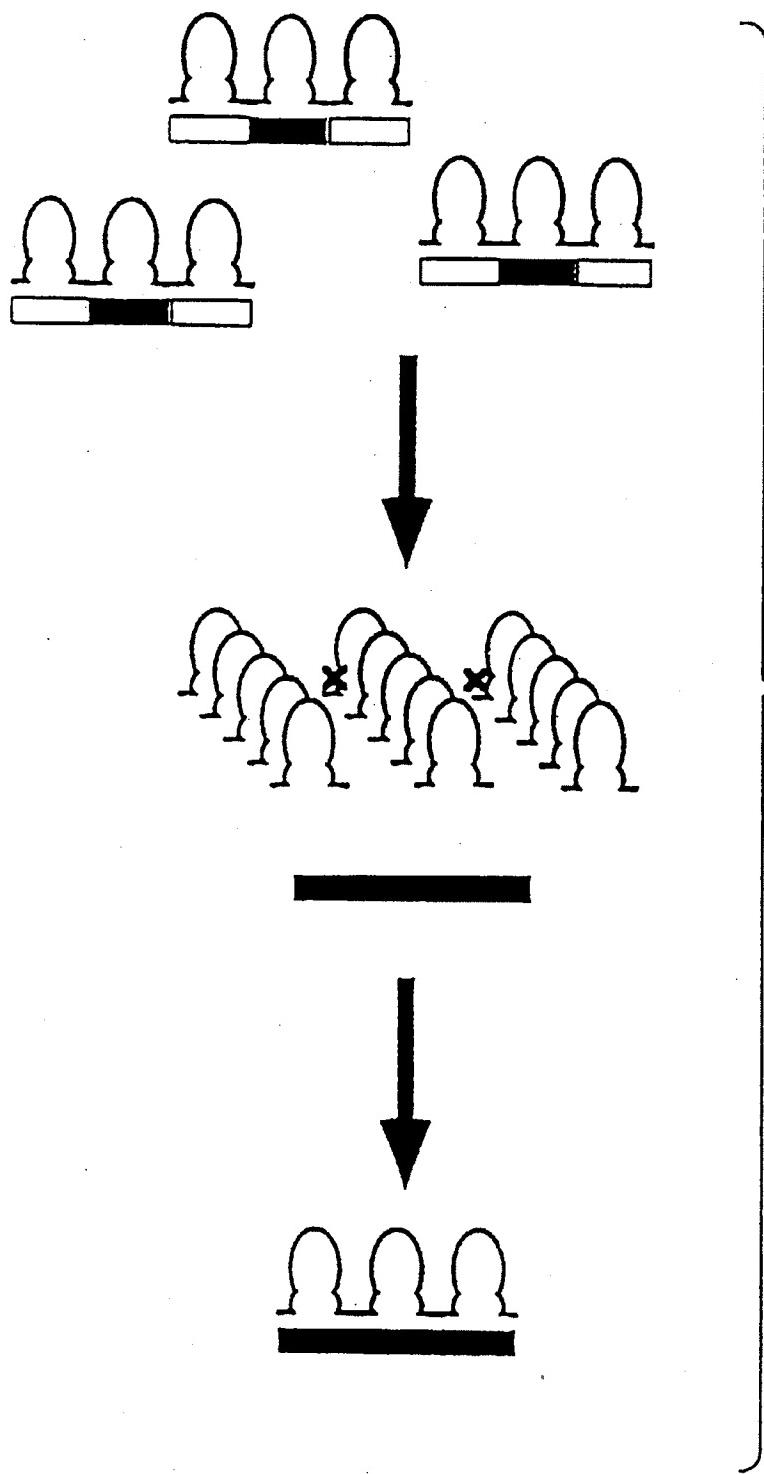


Fig. 11

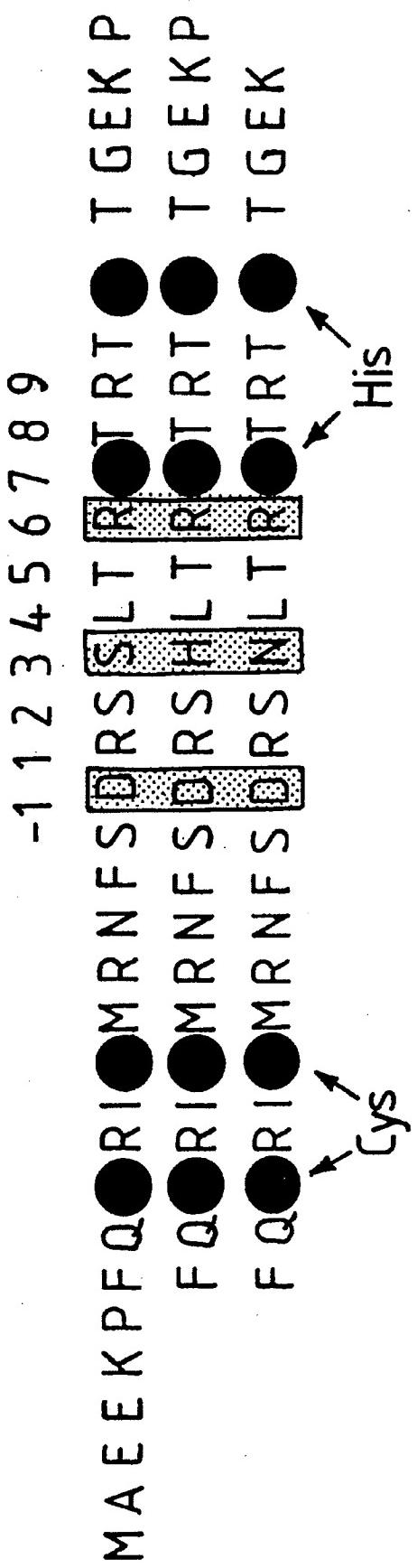


Fig. 12

BINDING PROTEINS FOR RECOGNITION OF DNA

This is a continuation of application No. 08/793,408, filed Jun. 2, 1997, which is a 371 of PCT/GB95/01949, filed Aug. 17, 1995.

FIELD OF THE INVENTION

This invention relates inter alia to methods of selecting and designing polypeptides comprising zinc finger binding motifs, polypeptides made by the method(s) of the invention and to various applications thereof.

BACKGROUND OF THE INVENTION

Selective gene expression is mediated via the interaction of protein transcription factors with specific nucleotide sequences within the regulatory region of the gene. The most widely used domain within protein transcription factors appears to be the zinc finger (Zf) motif. This is an independently folded zinc-containing mini-domain which is used in a modular repeating fashion to achieve sequence-specific recognition of DNA (Klug 1993 Gene 135, 83–92). The first zinc finger motif was identified in the Xenopus transcription factor TFIIIA (Miller et al., 1985 EMBO J. 4, 1609–1614). The structure of Zf proteins has been determined by NMR studies (Lee et al., 1989 Science 245, 635–637) and crystallography (Pavletich & Pabo, 1991 Science 252, 809–812).

The manner in which DNA-binding protein domains are able to discriminate between different DNA sequences is an important question in understanding crucial processes such as the control of gene expression in differentiation and development. The zinc finger motif has been studied extensively, with a view to providing, some insight into this problem, owing to its remarkable prevalence in the eukaryotic genome, and its important role in proteins which control gene expression in Drosophila (e.g., Harrison & Travers 1990 EMBO J. 9, 207–216), the mouse (Christy et al., 1988 Proc. Natl. Acad. Sci. USA 85, 7857–7861) and humans (Kinzel et al., 1988 Nature (London) 332, 371).

Most sequence-specific DNA-binding proteins bind to the DNA double helix by inserting an α -helix into the major groove (Pabo & Sauer 1992 Annu. Rev. Biochem. 61, 1053–1095; Harrison 1991 Nature (London) 353, 715–719; and Klug 1993 Gene 135, 83–92). Sequence specificity results from the geometrical and chemical complementarity between the amino acid side chains of the α -helix and the accessible groups exposed on the edges of base-pairs. In addition to this direct reading of the DNA sequence, interactions with the DNA backbone stabilise the complex and are sensitive to the conformation of the nucleic acid, which in turn depends on the base sequence (Dickerson & Drew 1981 J. Mol. Biol. 149, 761–786). A priori, a simple set of rules might suffice to explain the specific association of protein and DNA in all complexes, based on the possibility that certain amino acid side chains have preferences for particular base-pairs. However, crystal structures of protein-DNA complexes have shown that proteins can be idiosyncratic in their mode of DNA recognition, at least partly because they may use alternative geometries to present their sensory α -helices to DNA, allowing a variety of different base contacts to be made by a single amino acid and vice versa (Matthews 1988 Nature (London) 335, 294–295).

Mutagenesis of Zf proteins has confirmed modularity of the domains. Site directed mutagenesis has been used to change key Zf residues, identified through sequence homology alignment, and from the structural data, resulting in

altered specificity of Zf domain (Nardelli et al., 1992 NAR 26, 4137–4144). The authors suggested that although design of novel binding specificities would be desirable, design would need to take into account sequence and structural data. They state “there is no prospect of achieving a zinc finger recognition code”.

Despite this, many groups have been trying to work towards such a code, although only limited rules have so far been proposed. For example, Desjarlais et al., (1992b PNAS 89, 7345–7349) used systematic mutation of two of the three contact residues (based on consensus sequences) in finger two of the polypeptide Sp1 to suggest that a limited degenerate code might exist. Subsequently the authors used this to design three Zf proteins with different binding specificities and affinities (Desjarlais & Berg, 1993 PNAS 90, 2250–2260). They state that the design of Zf proteins with predictable specificities and affinities “may not always be straightforward”.

We believe the zinc finger of the TFIIIA class to be a good candidate for deriving a set of more generally applicable specificity rules owing to its great simplicity of structure and interaction with DNA. The zinc finger is an independently folding domain which uses a zinc ion to stabilise the packing of an antiparallel β -sheet against an α -helix (Miller et al., 1985 EMBO J. 4, 1609–1614; Berg 1988 Proc. Natl. Acad. Sci. USA 85, 99–102; and Lee et al., 1989 Science 245, 635–637). The crystal structures of zinc finger-DNA complexes show a semiconserved pattern of interactions in which 3 amino acids from the α -helix contact 3 adjacent bases (a triplet) in DNA (Pavletich & Pabo 1991 Science 252, 809–817; Fairall et al., 1993 Nature (London) 366, 483–487; and Pavletich & Pabo 1993 Science 261, 1701–1707). Thus the mode of DNA recognition is principally a one-to-one interaction between amino acids and bases. Because zinc fingers function as independent modules (Miller et al., 1985 EMBO J. 4, 1609–1614; Kiug & Rhodes 1987 Trends Biochem. Sci. 12, 464–469), it should be possible for fingers with different triplet specificities to be combined to give specific recognition of longer DNA sequences. Each finger is folded so that three amino acids are presented for binding to the DNA target sequence, although binding may be directly through only two of these positions. In the case of Zif268 for example, the protein is made up of three fingers which contact a 9 base pair contiguous sequence of target DNA. A linker sequence is found between fingers which appears to make no direct contact with the nucleic acid.

Protein engineering experiments have shown that it is possible to alter rationally the DNA-binding characteristics of individual zinc fingers when one or more of the α -helical positions is varied in a number of proteins (Nardelli et al., 1991 Nature (London) 349, 175–178; Nardelli et al., 1992 Nucleic Acids Res. 20, 4137–4144; and Desjarlais & Berg 1992a Proteins 13, 272). It has already been possible to propose some principles relating amino acids on the α -helix to corresponding bases in the bound DNA sequence (Desjarlais & Berg 1992b Proc. Natl. Acad. Sci. USA 89, 7345–7349). However in this approach the altered positions on the α -helix are prejudged, making it possible to overlook the role of positions which are not currently considered important; and secondly, owing to the importance of context, concomitant alterations are sometimes required to affect specificity (Desjarlais & Berg 1992b), so that a significant correlation between an amino acid and base may be misconstrued.

To investigate binding of mutant Zf proteins, Thiesen and Bach (1991 FEBS 283, 23–26) mutated Zf fingers and

studied their binding to randomised oligonucleotides, using electrophoretic mobility shift assays. Subsequent use of phage display technology has permitted the expression of random libraries of Zf mutant proteins on the surface of bacteriophage. The three Zf domains of Zif268, with 4 positions within finger one randomised, have been displayed on the surface of filamentous phage by Rebar and Pabo (1994 Science 263, 671-673). The library was then subjected to rounds of affinity selection by binding to target DNA oligonucleotide sequences in order to obtain Zf proteins with new binding specificities. Randomised mutagenesis (at the same positions as those selected by Rebar & Pabo) of finger 1 of Zif 268 with phage display has also been used by Jamieson et al., (1994 Biochemistry 33, 5689-5695) to create novel binding specificity and affinity.

More recently Wu et al. (1995 Proc. Natl. Acad. Sci. USA 92, 344-348) have made three libraries, each of a different finger from Zif268, and each having six or seven α -helical positions randomised. Six triplets were used in selections but did not return fingers with any sequence biases; and when the three triplets of the Zif268 binding site were individually used as controls, the vast majority of selected fingers did not resemble the sequences of the wild-type Zif268 fingers and, though capable of tight binding to their target sites in vitro, were usually not able to discriminate strongly against different triplets. The authors interpret the results as evidence against the existence of a code.

In summary, it is known that Zf protein motifs are widespread in DNA binding proteins and that binding is via three key amino acids, each one contacting a single base pair in the target DNA sequence. Motifs are modular and may be linked together to form a set of fingers which recognise a contiguous DNA sequence (e.g. a three fingered protein will recognise a 9mer etc). The key residues involved in DNA binding have been identified through sequence data and from structural information. Directed and random mutagenesis has confirmed the role of these amino acids in determining specificity and affinity. Phage display has been used to screen for new binding specificities of random mutants of fingers. A recognition code, to aid design of new finger specificities, has been worked towards although it has been suggested that specificity may be difficult to predict.

SUMMARY OF THE INVENTION

In a first aspect the invention provides a library of DNA sequences, each sequence encoding at least one zinc finger binding motif for display on a viral particle, the sequences coding for zinc finger binding motifs having random allocation of amino acids at positions -1, +2, +3, +6 and at least at one of positions +1, +5 and +8.

A zinc finger binding motif is the α -helical structural motif found in zinc finger binding proteins, well known to those skilled in the art. The above numbering is based on the first amino acid in the α -helix of the zinc finger binding motif being position +1. It will be apparent to those skilled in the art that the amino acid residue at position -1 does not, strictly speaking, form part of the α -helix of the zinc binding finger motif. Nevertheless, the residue at -1 is shown to be very important functionally and is therefore considered as part of the binding motif α -helix for the purposes of the present invention.

The sequences may code for zinc finger binding motifs having random allocation at all of positions +1, +5 and +8. The sequences may also be randomised at other positions (e.g. at position +9, although it is generally preferred to retain an arginine or a lysine residue at this position).

Further, whilst allocation of amino acids at the designated "random" positions may be genuinely random, it is preferred to avoid a hydrophobic residue (Phe, Trp or Tyr) or a cysteine residue at such positions.

Preferably the zinc finger binding motif is present within the context of other amino acids (which may be present in zinc finger proteins), so as to form a zinc finger (which includes an antiparallel β -sheet). Further, the zinc finger is preferably displayed as part of a zinc finger polypeptide, which polypeptide comprises a plurality of zinc fingers joined by an intervening linker peptide. Typically the library of sequences is such that the zinc finger polypeptide will comprise two or more zinc fingers of defined amino acid sequence (generally the wild type sequence) and one zinc finger having a zinc finger binding motif randomised in the manner defined above. It is preferred that the randomised finger of the polypeptide is positioned between the two or more fingers having defined sequence. The defined fingers will establish the "phase" of binding of the polypeptide to DNA, which helps to increase the binding specificity of the randomised finger.

Preferably the sequences encode the randomised binding motif of the middle finger of the Zif268 polypeptide. Conveniently, the sequences also encode those amino acids N-terminal and C-terminal of the middle finger in wild type Zif268, which encode the first and third zinc fingers respectively. In a particular embodiment, the sequence encodes the whole of the Zif268 polypeptide. Those skilled in the art will appreciate that alterations may also be made to the sequence of the linker peptide and/or the β -sheet of the zinc finger polypeptide.

In a further aspect, the invention provides a library of DNA sequences, each sequence encoding the zinc finger binding motif of at least a middle finger of a zinc finger binding polypeptide for display on a viral particle, the sequences coding for the binding motif having random allocation of amino acids at positions -1, +2, +3 and +6. Conveniently, the zinc finger polypeptide will be Zif268.

Typically, the sequences of either library are such that the zinc finger binding domain can be cloned as a fusion with the minor coat protein (pIII) of bacteriophage fd. Conveniently, the encoded polypeptide includes the tripeptide sequence Met-Ala-Glu as the N terminal of the zinc finger domain, which is known to allow expression and display using the bacteriophage fd system. Desirably the library comprises 10^6 or more different sequences (ideally, as many as is practicable).

In another aspect the invention provides a method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising screening each of a plurality of zinc finger binding motifs against at least an effective portion of the target DNA sequence, and selecting those motifs which bind to the target DNA sequence. An effective portion of the target DNA sequence is a sufficient length of DNA to allow binding of the zinc binding motif to the DNA. This is the minimum sequence information (concerning the target DNA sequence) that is required. Desirably at least two, preferably three or more, rounds of screening are performed.

The invention also provides a method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising comparing the binding of each of a plurality of zinc finger binding motifs to one or more DNA triplets, and selecting those motifs exhibiting preferable binding characteristics. Preferably the method defined immediately above is preceded by a screening step according to the method defined in the previous paragraph.

It is thus preferred that there is a two-step selection procedure: the first step comprising screening each of a plurality of zinc finger binding motifs (typically in the form of a display library), mainly or wholly on the basis of affinity for the target sequence; the second step comprising comparing binding characteristics of those motifs selected by the initial screening step, and selecting those having preferable binding characteristics for a particular DNA triplet.

Where the plurality of zinc finger binding motifs is screened against a single DNA triplet, it is preferred that the triplet is represented in the target DNA sequence at the appropriate position. However, it is also desirable to compare the binding of the plurality of zinc binding motifs to one or more DNA triplets not represented in the target DNA sequence (e.g. differing by just one of the three base pairs) in order to compare the specificity of binding of the various binding motifs. The plurality of zinc finger binding motifs may be screened against all 64 possible permutations of 3 DNA bases.

Once suitable zinc finger binding motifs have been identified and obtained, they will advantageously be combined in a single zinc finger polypeptide. Typically this will be accomplished by use of recombinant DNA technology; conveniently a phage display system may be used.

In another aspect, the invention provides a DNA library consisting of 64 sequences, each sequence comprising a different one of the 64 possible permutations of three DNA bases in a form suitable for use in the selection method defined above. Desirably the sequences are associated, or capable of being associated, with separation means. Advantageously, the separation means is selected from one of the following: microtitre plate; magnetic beads; or affinity chromatography column. Conveniently the sequences are biotinylated. Preferably the sequences are contained within 12 mini-libraries, as explained elsewhere.

In a further aspect the invention provides a zinc finger polypeptide designed by one or both of the methods defined above. Preferably the zinc finger polypeptide designed by the method comprises a combination of a plurality of zinc fingers (adjacent zinc fingers being joined by an intervening linker peptide), each finger comprising a zinc finger binding motif. Desirably, each zinc finger binding motif in the zinc finger polypeptide has been selected for preferable binding characteristics by the method defined above. The intervening linker peptide may be the same between each adjacent zinc finger or, alternatively, the same zinc finger polypeptide may contain a number of different linker peptides. The intervening linker peptide may be one that is present in naturally-occurring zinc finger polypeptides or may be an artificial sequence. In particular, the sequence of the intervening linker peptide may be varied, for example, to optimise binding of the zinc finger polypeptide to the target sequence.

Where the zinc finger polypeptide comprises a plurality of zinc binding motifs, it is preferred that each motif binds to those DNA triplets which represent contiguous or substantially contiguous DNA in the sequence of interest. Where several candidate binding motifs or candidate combinations of motifs exist, these may be screened against the actual target sequence to determine the optimum composition of the polypeptide. Competitor DNA may be included in the screening assay for comparison, as described below.

The non-specific component of all protein-DNA interactions, which includes contacts to the sugar-phosphate backbone as well as ambiguous contacts to base-pairs, is a considerable driving force towards complex formation and

can result in the selection of DNA-binding proteins with reasonable affinity but without specificity for a given DNA sequence. Therefore, in order to minimise these non-specific interactions when designing a polypeptide, selections should preferably be performed with low concentrations of specific binding site in a background of competitor DNA, and binding should desirably take place in solution to avoid local concentration effects and the avidity of multivalent phage for ligands immobilised on solid surfaces.

As a safeguard against spurious selections, the specificity of individual phage should be determined following the final round of selection. Instead of testing for binding to a small number of binding sites, it would be desirable to screen all possible DNA sequences.

It has now been shown possible by the present inventors (below) to design a truly modular zinc binding polypeptide, wherein the zinc binding motif of each zinc binding finger is selected on the basis of its affinity for a particular triplet. Accordingly, it should be well within the capability of one

of normal skill in the art to design a zinc finger polypeptide capable of binding to any desired target DNA sequence simply by considering the sequence of triplets present in the target DNA and combining in the appropriate order zinc fingers comprising zinc finger binding motifs having the necessary binding characteristics to bind thereto. The greater

the length of known sequence of the target DNA, the greater the number of zinc finger binding motifs that can be included in the zinc finger polypeptide. For example, if the known sequence is only 9 bases long then three zinc finger binding motifs can be included in the polypeptide. If the known sequence is 27 bases long then, in theory, up to nine binding motifs could be included in the polypeptide. The longer the target DNA sequence, the lower the probability of its occurrence in any given portion of DNA.

Moreover, those motifs selected for inclusion in the polypeptide could be artificially modified (e.g. by directed mutagenesis) in order to optimise further their binding characteristics. Alternatively (or additionally) the length and amino acid sequence of the linker peptide joining adjacent zinc binding fingers could be varied, as outlined above. This may have the effect of altering the position of the zinc finger binding motif relative to the DNA sequence of interest, and thereby exert a further influence on binding characteristics.

Generally, it will be preferred to select those motifs having high affinity and high specificity for the target triplet.

In a further aspect, the invention provides a kit for making a zinc finger polypeptide for binding to a nucleic acid sequence of interest, comprising: a library of DNA sequences encoding zinc finger binding motifs of known binding characteristics in a form suitable for cloning into a vector; a vector molecule suitable for accepting one or more sequences from the library; and instructions for use.

Preferably the vector is capable of directing the expression of the cloned sequences as a single zinc finger polypeptide. In particular it is preferred that the vector is capable of directing the expression of the cloned sequences as a single zinc finger polypeptide displayed on the surface of a viral particle, typically of the sort of viral display particle which are known to those skilled in the art. The DNA sequences are preferably in such a form that the expressed polypeptides are capable of self-assembling into a number of zinc finger polypeptides.

It will be apparent that the kit defined above will be of particular use in designing a zinc finger polypeptide comprising a plurality of zinc finger binding motifs, the binding characteristics of which are already known. In another

aspect the invention provides a kit for use when zinc finger binding motifs with suitable binding characteristics have not yet been identified, such that the invention provides a kit for making a zinc finger polypeptide for binding to a nucleic acid sequence of interest, comprising: a library of DNA sequences, each encoding a zinc finger binding motif in a form suitable for screening and/or selecting according to the methods defined above; and instructions for use.

Advantageously, the library of DNA sequences in the kit will be a library in accordance with the first aspect of the invention. Conveniently, the kit may also comprise a library of 64 DNA sequences, each sequence comprising a different one of the 64 possible permutations of three DNA bases, in a form suitable for use in the selection method defined previously. Typically, the 64 sequences are present in 12 separate mini-libraries, each mini-library having one position in the relevant triplet fixed and two positions randomised. Preferably, the kit will also comprise appropriate buffer solutions, and/or reagents for use in the detection of bound zinc fingers. The kit may also usefully include a vector suitable for accepting one or more sequences selected from the library of DNA sequences encoding zinc finger binding motifs.

In a preferred embodiment, the present teaching will be used for isolating the genes for the middle zinc fingers which, having been previously selected by one of the 64 triplets, are thought to have specific DNA binding activity. The mixture of genes specifying fingers which bind to a given triplet will be amplified by PCR using three sets of primers. The sets will have unique restriction sites, which will define the assembly of zinc fingers into three finger polypeptides. The appropriate reagents are preferably provided in kit form.

For instance, the first set of primers might have SfiI and AgeI sites, the second set AgeI and EagI sites and third set EagI and NotI sites. It will be noted that the "first" site will preferably be SfiI, and the "last" site NotI, so as to facilitate cloning into the SfiI and NotI sites of the phage vector. To assemble a library of three finger proteins which recognise the sequence AAAGGGGG, the fingers selected by the triplet GGG are amplified using the first two sets of primers and ligated to the fingers selected by the triplet AAA amplified using the third set of primers. The combinatorial library is cloned on the surface of phage and a nine base-pair site can be used to select the best combination of fingers en bloc.

The genes for fingers which bind to each of the 64 triplets can be amplified by each set of primers and cut using the appropriate restriction enzymes. These building blocks for three-finger proteins can be sold as components of a kit for use as described above. The same could be done for the library amplified with different primers so that 4- or 5- finger proteins could be built.

Additionally a large (pre-assembled) library of all combinations of the fingers selected by all triplets can also be developed for single-step selection of DNA-binding proteins using 9 bp, or much longer, DNA fragments. For this particular application, which will require very large libraries of novel 3-finger proteins, it may be preferable to use methods of selection other than phage display; for example stalled polysomes (developed by Affimax) where protein and mRNA become linked.

In a further aspect the invention provides a method of altering the expression of a gene of interest in a target cell, comprising: determining (if necessary) at least part of the DNA sequence of the structural region and/or a regulatory

region of the gene of interest; designing a zinc finger polypeptide to bind to the DNA of known sequence, and causing said zinc finger polypeptide to be present in the target cell, (preferably in the nucleus thereof). (It will be apparent that the DNA sequence need not be determined if it is already known.)

The regulatory region could be quite remote from the structural region of the gene of interest (e.g. a distant enhancer sequence or similar). Preferably the zinc finger polypeptide is designed by one or both of the methods of the invention defined above.

Binding of the zinc finger polypeptide to the target sequence may result in increased or reduced expression of the gene of interest depending, for example, on the nature of the target sequence (e.g. structural or regulatory) to which the polypeptide binds.

In addition, the zinc finger polypeptide may advantageously comprise functional domains from other proteins (e.g. catalytic domains from restriction enzymes, recombinases, replicases, integrases and the like) or even "synthetic" effector domains. The polypeptide may also comprise activation or processing signals, such as nuclear localisation signals. These are of particular usefulness in targeting the polypeptide to the nucleus of the cell in order to enhance the binding of the polypeptide to an intranuclear target (such as genomic DNA). A particular example of such a localisation signal is that from the large T antigen of SV40. Such other functional domains/signals and the like are conveniently present as a fusion with the zinc finger polypeptide. Other desirable fusion partners comprise immunoglobulins or fragments thereof (e.g. Fab, scFv) having binding activity.

The zinc finger polypeptide may be synthesised in situ in the cell as a result of delivery to the cell of DNA directing expression of the polypeptide. Methods of facilitating delivery of DNA are well-known to those skilled in the art and include, for example, recombinant viral vectors (e.g. retroviruses, adenoviruses), liposomes and the like. Alternatively, the zinc finger polypeptide could be made outside the cell and then delivered thereto. Delivery could be facilitated by incorporating the polypeptide into liposomes etc. or by attaching the polypeptide to a targeting moiety (such as the binding portion of an antibody or hormone molecule). Indeed, one significant advantage of zinc finger proteins over oligonucleotides or protein-nucleic acids (PNAs) in controlling gene expression, would be the vector-free delivery of protein to target cells. Unlike the above, many examples of soluble proteins entering cells are known, including antibodies to cell surface receptors. The present inventors are currently carrying out fusions of anti-bcr-abl fingers (see example 3 below) to a single-chain (sc) Fv fragment capable of recognising NIP (4-hydroxy-5-iodo-3-nitrophenyl acetyl). Mouse transferrin conjugated with NIP will be used to deliver the fingers to mouse cells via the mouse transferrin receptor.

Media (e.g. microtitre wells, resins etc.) coated with NIP can also be used as solid supports for zinc fingers fused to anti-NIP scFvs, for applications requiring immobilised zinc fingers (e.g. the purification of specific nucleic acids).

In a particular embodiment, the invention provides a method of inhibiting cell division by causing the presence in a cell of a zinc finger polypeptide which inhibits the expression of a gene enabling the cell to divide.

In a specific embodiment, the invention provides a method of treating a cancer, comprising delivering to a patient, or causing to be present therein, a zinc finger

polypeptide which inhibits the expression of a gene enabling the cancer cells to divide. The target could be, for example, an oncogene or a normal gene which is overexpressed in the cancer cells.

To the best knowledge of the inventors, design of a zinc finger polypeptide and its successful use in modulation of gene expression (as described below) has never previously been demonstrated. This breakthrough presents numerous possibilities. In particular, zinc finger polypeptides could be designed for therapeutic and/or prophylactic use in regulating the expression of disease-associated genes. For example, zinc finger polypeptides could be used to inhibit the expression of foreign genes (e.g. the genes of bacterial or viral pathogens) in man or animals, or to modify the expression of mutated host genes (such as oncogenes).

The invention therefore provides a zinc finger polypeptide capable of inhibiting the expression of a disease-associated gene. Typically the zinc finger polypeptide will not be a naturally-occurring polypeptide but will be specifically designed to inhibit the expression of the disease-associated gene. Conveniently the polypeptide will be designed by one or both of the methods of the invention defined above. Advantageously the disease-associated gene will be an oncogene, typically the BCR-ABL fusion oncogene or a ras oncogene. In a particular embodiment the invention provides a zinc finger polypeptide designed to bind to the DNA sequence GCAGAAAGCC and capable of inhibiting the expression of the BCR-ABL fusion oncogene.

In yet another aspect the invention provides a method of modifying a nucleic acid sequence of interest present in a sample mixture by binding thereto a zinc finger polypeptide, comprising contacting the sample mixture with a zinc finger polypeptide having affinity for at least a portion of the sequence of interest, so as to allow the zinc finger polypeptide to bind specifically to the sequence of interest.

The term "modifying" as used herein is intended to mean that the sequence is considered modified simply by the binding of the zinc finger polypeptide. It is not intended to suggest that the sequence of nucleotides is changed, although such changes (and others) could ensue following binding of the zinc finger polypeptide to the nucleic acid of interest. Conveniently the nucleic acid sequence is DNA.

Modification of the nucleic acid of interest (in the sense of binding thereto by a zinc finger polypeptide) could be detected in any of a number of methods (e.g. gel mobility shift assays, use of labelled zinc finger polypeptides—labels could include radioactive, fluorescent, enzyme or biotin/streptavidin labels).

Modification of the nucleic acid sequence of interest (and detection thereof) may be all that is required (e.g. in diagnosis of disease). Desirably however, further processing of the sample is performed. Conveniently the zinc finger polypeptide (and nucleic acid sequences specifically bound thereto) are separated from the rest of the sample. Advantageously the zinc finger polypeptide is bound to a solid phase support, to facilitate such separation. For example, the zinc finger polypeptide may be present in an acrylamide or agarose gel matrix or, more preferably, is immobilised on the surface of a membrane or in the wells of a microtitre plate.

Possible uses of suitably designed zinc finger polypeptides are:

- Therapy (e.g. targeting to double stranded DNA)
- Diagnosis (e.g. detecting mutations in gene sequences: the present work has shown that "tailor made" zinc finger polypeptides can distinguish DNA sequences differing by one base pair).

c) DNA purification (the zinc finger polypeptide could be used to purify restriction fragments from solution, or to visualise DNA fragments on a gel [for example, where the polypeptide is linked to an appropriate fusion partner, or is detected by probing with an antibody]).

In addition, zinc finger polypeptides could even be targeted to other nucleic acids such as ss or ds RNA (e.g. self-complementary RNA such as is present in many RNA molecules) or to RNA-DNA hybrids, which would present another possible mechanism of affecting cellular events at the molecular level.

In Example 1 the inventors describe and successfully demonstrate the use of the phage display technique to construct and screen a random zinc finger binding motif library, using a defined oligonucleotide target sequence.

In Example 2 is disclosed the analysis of zinc finger binding motif sequences selected by the screening procedure of Example 1, the DNA-specificity of the motifs being studied by binding to a mini-library of randomised DNA target sequences to reveal a pattern of acceptable bases at each position in the target triplet—a "binding site signature".

In Example 3, the findings of the first two sections are used to select and modify rationally a zinc finger binding polypeptide in order to bind to a particular DNA target with high affinity: it is convincingly shown that the peptide binds to the target sequence and can modify gene expression in cells cultured *in vitro*.

Example 4 describes the development of an alternative zinc finger binding motif library.

Example 5 describes the design of a zinc finger binding polypeptide which binds to a DNA sequence of special clinical significance.

The invention will now be further described by way of example and with reference to the accompanying drawings, of which:

FIG. 1 is a schematic representation of affinity purification of phage particles displaying zinc finger binding motifs fused to phage coat proteins;

FIG. 2 shows three zinc fingers (Seq ID No. 2) used in the phage display library;

FIG. 3 shows the DNA sequences of three oligonucleotides (Seq ID No.s 3-8) used in the affinity purification of phage display particles;

FIG. 4 is a "checker board" of binding site signatures determined for various zinc finger binding motifs (Seq ID No.s 19-51);

FIGS. 5A-5F show three of fractional saturation against concentration of DNA (nM) for various binding motifs and target DNA triplets;

FIG. 6 shows the nucleotide sequence of the fusion between BCR and ABL sequences in p190 cDNA (Seq ID No. 9) and the corresponding exon boundaries in the BCR and ABL genes (Seq ID No.s 10-11);

FIG. 7 shows the amino acid sequences of various zinc finger binding motifs (Seq ID No.s 12-17) designed to test for binding to the BCR/ABL fusion;

FIG. 8 is a graph of peptide binding (as measured by $A_{450-460}$ nm) against DNA concentration (μM) of target or control DNA sequences;

FIG. 9 is a graph showing percentage viability against time for various transfected cells;

FIGS. 10A-10C and 11 illustrate schematically different methods of designing zinc finger binding polypeptides; and

FIG. 12 shows the amino acid sequence of zinc fingers in a polypeptide (Seq ID No. 18) designed to bind to a particular DNA sequence (a ras oncogene).

EXAMPLE 1

In this example the inventors have used a screening technique to study sequence-specific DNA recognition by

zinc finger binding motifs. The example describes how a library of zinc finger binding motifs displayed on the surface of bacteriophage enables selection of fingers capable of binding to given DNA triplets. The amino acid sequences of selected fingers which bind the same triplet were compared to examine how sequence-specific DNA recognition occurs. The results can be rationalised in terms of coded interactions between zinc fingers and DNA, involving base contacts from a few α -helical positions.

An alternative to the rational but biased design of proteins with new specificities, is the isolation of desirable mutants from a large pool. A powerful method of selecting such proteins is the cloning of peptides (Smith 1985 *Science* 228, 1315–1317), or protein domains (McCafferty et al., 1990 *Nature* (London) 348, 552–554; Bass et al., 1990 *Proteins* 8, 309–314), as fusions to the minor coat protein (pIII) of bacteriophage fd, which leads to their expression on the tip of the capsid. Phage displaying the peptides of interest can then be affinity purified and amplified for use in further rounds of selection and for DNA sequencing of the cloned gene. The inventors applied this technology to the study of zinc finger-DNA interactions after demonstrating that functional zinc finger proteins can be displayed on the surface of fd phage, and that the engineered phage can be captured on a solid support coated with specific DNA. A phage display library was created comprising variants of the middle finger from the DNA binding domain of Zif268 (a mouse transcription factor containing 3 zinc fingers—Christy et al., 1988). DNA of fixed sequence was used to purify phage from this library over several rounds of selection, returning a number of different but related zinc fingers which bind the given DNA. By comparing similarities in the amino acid sequences of functionally equivalent fingers we deduce the likely mode of interaction of these fingers with DNA. Remarkably, it would appear that many base contacts can occur from three primary positions on the α -helix of the zinc finger, correlating (in hindsight) with the implications of the crystal structure of Zif268 bound to DNA (Pavletich & Pabo 1991). The ability to select or design zinc fingers with desired specificity means that DNA binding proteins containing zinc fingers can now be “made-to-measure”.

MATERIALS AND METHODS

Construction and cloning of genes. The gene for the first three fingers (residues 3–101) of Transcription Factor IIIA (TFIIIA) was amplified by PCR from the cDNA clone of TFIIIA using forward and backward primers which contain restriction sites for NotI and SfiI respectively. The gene for the Zif268 fingers (residues 333–420) was assembled from 8 overlapping synthetic oligonucleotides, giving SfiI and NotI overhangs. The genes for fingers of the phage library were synthesised from 4 oligonucleotides by directional end to end ligation using 3 short complementary linkers, and amplified by PCR from the single strand using forward and backward primers which contained sites for NotI and SfiI respectively. Backward PCR primers in addition introduced Met-Ala-Glu as the first three amino acids of the zinc finger peptides, and these were followed by the residues of the wild type or library fingers as discussed in the text. Cloning overhangs were produced by digestion with SfiI and NotI where necessary. Fragments were ligated to 1 μ g similarly prepared Fd-Tet-SN vector. This is a derivative of fd-tet-DOG1 (Hoogenboom et al., 1991 *Nucleic Acids Res.* 19, 4133–4137) in which a section of the pELB leader and a restriction site for the enzyme SfiI (underlined) have been added by site-directed mutagenesis using the oligonucleotide (Seq ID No. 1):

5' CTCCTGCAGTTGGACCTGTGCCATGGCCG
GCTGGCCGCATAGAACAACTAAAGC 3'

which anneals in the region of the polylinker, (L. Jespers, personal communication). Electrocompetent DH5 α cells were transformed with recombinant vector in 200 ng aliquots, grown for 1 hour in 2xTY medium with 1% glucose, and plated on TYE containing 15 μ g/ml tetracycline and 1% glucose.

FIG. 2 shows the amino acid sequence (Seq ID No. 2) of the three zinc fingers from Zif268 used in the phage display library. The top and bottom rows represent the sequence of the first and third fingers respectively. The middle row represents the sequence of the middle finger. The randomised positions in the α -helix of the middle finger have residues marked ‘X’. The amino acid positions are numbered relative to the first helical residue (position 1). For amino acids at positions -1 to +8, excluding the conserved Leu and His, codons are equal mixtures of (G,A,C)NN: T in the first base position is omitted in order to avoid stop codons, but this has the unfortunate effect that the codons for Trp, Phe, Tyr and Cys are not represented. Position +9 is specified by the codon A(G,A)G, allowing either Arg or Lys. Residues of the hydrophobic core are circled, whereas the zinc ligands are written as white letters on black circles. The positions forming the β -sheets and the α -helix of the zinc fingers are marked below the sequence.

Phage selection. Colonies were transferred from plates to 200 ml 2xTY/Zn/Tet (2xTY containing 50 μ M Zn(CH₃COO)₂ and 15 μ g/ml tetracycline) and grown overnight. Phage were purified from the culture supernatant by two rounds of precipitation using 0.2 volumes of 20% PEG/2.5M NaCl containing 50 μ M Zn(CH₃COO)₂, and resuspended in zinc finger phage buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM MgCl₂ and 50 μ M Zn(CH₃COO)₂). Streptavidin-coated paramagnetic beads (Dynal) were washed in zinc finger phage buffer and blocked for 1 hour at room temperature with the same buffer made up to 6% in fat-free dried milk (Marvel). Selection of phage was over three rounds: in the first round, beads (1 mg) were saturated with biotinylated oligonucleotide (~80 nM) and then washed prior to phage binding, but in the second and third rounds 1.7 nM oligonucleotide and 5 μ g poly dGC (Sigma) were added to the beads with the phage. Binding reactions (1.5 ml) for 1 hour at 15° C. were in zinc finger phage buffer made up to 2% in fat-free dried milk (Marvel) and 1% in Tween 20, and typically contained 5 \times 10¹¹ phage. Beads were washed 15 times with 1 ml of the same buffer. Phage were eluted by shaking in 0.1M triethylamine for 5 min and neutralised with an equal volume of 1M Tris pH7.4. Log phase *E. coli* TG1 in 2xTY were infected with eluted phage for 30 min at 37° C. and plated as described above. Phage titres were determined by plating serial dilutions of the infected bacteria.

The phage selection procedure, based on affinity purification, is illustrated schematically in FIG. 1: zinc fingers (A) are expressed on the surface of fd phage(B) as fusions to the minor coat protein (C). The third finger is mainly obscured by the DNA helix. Zinc finger phage are bound to 5'-biotinylated DNA oligonucleotide [D] attached to streptavidin-coated paramagnetic beads [E], and captured using a magnet [F]. (Figure adapted from Dynal A. S. and also Marks et al. (1992 *J. Biol. Chem.* 267, 16007–16105).

FIG. 3 shows sequences (Seq ID No.s 3–8) of DNA oligonucleotides used to purify (i) phage displaying the first three fingers of TFIIIA, (ii) phage displaying the three fingers of Zif268, and (iii) zinc finger phage from the phage

display library. The Zif268 consensus operator sequence used in the X-ray crystal structure (Pavletich & Pabo 1991 Science 252, 809-817) is highlighted in (ii), and in (iii) where "X" denotes a base change from the ideal operator in oligonucleotides used to purify phage with new specificities. Biotinylation of one strand is shown by a circled "B".

Sequencing of selected phage. Single colonies of transformants obtained after three rounds of selection as described, were grown overnight in 2×TY/Zn/Tet. Small aliquots of the cultures were stored in 15% glycerol at -20° C., to be used as an archive. Single-stranded DNA was prepared from phage in the culture supernatant and sequenced using the modified T7 RNA polymerase SEQUE-NASE™ 2.0 kit (U.S. Biochemical Corp.).

RESULTS AND DISCUSSION

Phage display of 3-finger DNA-Binding Domains from TFIII or Zif268. Prior to the construction of a phage display library, the inventors demonstrated that peptides containing three fully functional zinc fingers could be displayed on the surface of viable fd phage when cloned in the vector Fd-Tet-SN. In preliminary experiments, the inventors cloned as fusions to pIII firstly the three N-terminal fingers from TFIIIA (Ginsberg et al., 1984 Cell 39, 479-489), and secondly the three fingers from Zif-268 (Christy et al., 1988), for both of which the DNA binding sites are known. Peptide fused to the minor coat protein was detected in Western blots using an anti-pIII antibody (Stengele et al., 1990 J. Mol. Biol. 212, 143-149). Approximately 10-20% of total pIII in phage preparations was present as fusion protein.

Phage displaying either set of fingers were capable of binding to specific DNA oligonucleotides, indicating that zinc fingers were expressed and correctly folded in both instances. Paramagnetic beads coated with specific oligonucleotide were used as a medium on which to capture DNA-binding phage, and were consistently able to return between 100 and 500-fold more such phage, compared to free beads or beads coated with non-specific DNA. Alternatively, when phage displaying the three fingers of Zif268 were diluted 1:1.7×10³ with Fd-Tet-SN phage not bearing zinc fingers, and the mixture incubated with beads coated with Zif-268 operator DNA, one in three of the total phage eluted and transfected into *E. coli* were shown by colony hybridisation to carry the Zif268 gene, indicating an enrichment factor of over 500 for the zinc finger phage. Hence it is clear that zinc fingers displayed on fd phage are capable of preferential binding to DNA sequences with which they can form specific complexes, making possible the enrichment of wanted phage by factors of up to 500 in a single affinity purification step. Therefore, over multiple rounds of selection and amplification, very rare clones capable of sequence-specific DNA binding, can be selected from a large library.

A phage display library of zinc fingers from Zif268. The inventors have made a phage display library of the three fingers of Zif268 in which selected residues in the middle finger are randomised (FIG. 2), and have isolated phage bearing zinc fingers with desired specificity using a modified Zif268 operator sequence (Christy & Nathans 1989 Proc. Natl. Acad. Sci. USA 86, 8737-8741) in which the middle DNA triplet is altered to the sequence of interest (FIG. 3). In order to be able to study both the primary and secondary putative base recognition positions which are suggested by database analysis (Jacobs 1992 EMBO J. 11, 4507-4517), the inventors have designed the library of the middle finger so that, relative to the first residue in the α-helix (position +1), positions -1 to +8, but excluding the conserved Leu and His, can be any amino acid except Phe, Tyr, Trp and Cys

which occur only rarely at those positions (Jacobs 1993 Ph.D. thesis, University of Cambridge). In addition, the inventors have allowed position +9 (which might make an inter-finger contact with Ser at position -2 (Pavletich & Pabo 1991)) to be either Arg or Lys, the two most frequently occurring residues at that position.

The logic of this protocol, based upon the Zif268 crystal structure (Pavletich & Pabo 1991), is that the randomised finger is directed to the central triplet since the overall register of protein-DNA contacts is fixed by its two neighbours. This allows the examination of which amino acids in the randomised finger are the most important in forming specific complexes with DNA of known sequence. Since comprehensive variations are programmed in all the putative contact positions of the α-helix, it is possible to conduct an objective study of the importance of each position in DNA-binding (Jacobs 1992).

The size of the phage display library required, assuming full degeneracy of the 8 variable positions, is $(16^7 \times 2^1) = 5.4 \times 10^8$, but because of practical limitations in the efficiency of transformation with Fd-Tet-SN, the inventors were able to clone only 2.6×10^6 of these. The library used is therefore some two hundred times smaller than the theoretical size necessary to cover all the possible variations of the α-helix. Despite this shortfall, it has been possible to isolate phage which bind with high affinity and specificity to given DNA sequences, demonstrating the remarkable versatility of the zinc finger motif. Amino acid-base contacts in zinc finger-DNA complexes deduced from phage display selection. Of the 64 base triplets that could possibly form the binding site for variations of finger 2, the inventors have so far used 32 in attempts to isolate zinc finger phage as described. Results from these selections are shown in Table 1, which lists amino acid sequences (Seq ID No.s 52-118) of the variant α-helical regions from clones of library phage selected after 3 rounds of screening with variants of the Zif268 operator.

TABLE 1

			-123456789
a	CAG	1 9	RGDHLKSHIK RSDHLTHIR
b	TGA	3 1 (3)	QLAHLSTSHKR QSVLQLSHSR QKCHLTSHRK
c	GAA	2 1 1 2 1	QGGNLVRLHRL NGGNLGRHMK HRSNLLEHTTR KOSNLVSHQR IASNLLEHQH
d	GAT	1 1 1 1 1 1 1 1 1	DRSNLLEHTTR DOSNLLETHHR QOSHLVSHQR MCGNLGRHMK MGAHLVSHHR MCGNLQSHGR SHPNLNSHLK EPGMLTRHGR
e	GAC	4 1	DRSNLLEHTTR DHNALASHTR

TABLE 1-continued

<i>f</i>	GCC	2 7 1	RSS ₁ LTRHTR RGR ₁ LARHEK RRR ₁ LDRHQR
<i>g</i>	GTC	6 1	RSS ₁ LTRHTR RTS ₁ LSEHHR
<i>h</i>	GCA	1 2 2 1	SAG ₁ LVEHSHK GAQ ₁ LQRHLK TKA ₁ LASHMK EGG ₁ LARHER
<i>i</i>	GCT	1 1 1 1 1 1 1	RGS ₁ LGRHTR EKA ₁ LASHMK GAQ ₁ LQRHLK ERG ₁ LARHEK GRD ₁ LARHQK SGP ₁ LARHGR SRD ₁ LRRHNR
<i>j</i>	ACG	8 1	BRD ₁ LMNHIR RKD ₁ LYSHVR
<i>k</i>	ATG	8 1 1	BRD ₁ LMNHIR RGD ₁ LTSHER RVD ₁ LEKRR
<i>l</i>	GTA	1 1 (1) (2) 2 1 (1)	RDS ₁ LTRHTR ERT ₁ LSEHHR GAR ₁ LTRHQR EGG ₁ LARHER DRS ₁ LASHMR BRD ₁ LTRHSHK ERG ₁ LARHER
<i>m</i>	TTG	9 1	RGD ₁ LTSHER RAD ₁ LMYHKR
<i>n</i>	CCG	5 1 2 1 1	RQD ₁ LVGHHER RQS ₁ LVRHTR RAA ₁ LNRHVR RKD ₁ LVSHVVR BRD ₁ LMNHIR
<i>o</i>	GCG	1 3 1 3 1	RSD ₁ LKKHGK EGP ₁ LASHGR ARE ₁ LQRHTR MED ₁ LTHGK QSD ₁ LQRHHK
<i>p</i>	GTG	1 1 1 1 1 2 1	RLD ₁ LRTHIK RGD ₁ LTSHER RAD ₁ LMYHKR RVD ₁ LEKRR RDR ₁ LLNHIR RED ₁ LTHGK RSD ₁ LQRHHK

In Table 1, the amino acid sequences, aligned in the one letter code, are listed alongside the DNA oligonucleotides (*a* to *p*) used in their purification. The latter are denoted by the sequence of the central DNA triplet in the "bound" strand of the variant Zif268 operator. The amino acid positions are numbered relative to the first helical residue (position 1), and the three primary recognition positions are highlighted. The accompanying numbers indicate the independent occurrences of that clone in the sequenced population (5–10 colonies); where numbers are in parentheses, the clone(s) were detected in the penultimate round of selection but not in the final round. In addition to the DNA triplets shown here, others were also used in attempts to select zinc finger phage from the library, but most selected two clones, one having the α -helical sequence KASNLVSHIR (Seq ID No. 119), and the other having the sequence LRHNLETHMR (Seq ID No. 120). Those triplets were: ACT, AAA, TTT, CCT, CTT, TTC, AGT, CGA, CAT, AGA, AGC and AAT.

In general the inventors have been unable to select zinc fingers which bind specifically to triplets without a 5' or 3' guanine, all of which return the same limited set of phage after three rounds of selection (see). However for each of the 5 other triplets used to screen the library, a family of zinc finger phage is recovered. In these families is found a sequence bias in the randomised α -helix, which is interpreted as revealing the position and identity of amino acids used to contact the DNA. For instance: the middle fingers 10 from the 8 different clones selected with the triplet GAT (Table 1d) all have Asn at position +3 and Arg at position +6, just as does the first zinc finger of the Drosophila protein tramtrack in which they are seen making contacts to the same triplet in the cocrystal with specific DNA (Fairall et al., 15 1993). This indicates that the positional recurrence of a particular amino acid in functionally equivalent fingers is unlikely to be coincidental, but rather because it has a functional role. Thus using data collected from the phage display library (Table 1) it is possible to infer most of the 20 specific amino acid-DNA interactions. Remarkably, most of the results can be rationalised in terms of contacts from the three primary α -helical positions (-1, +3 and +6) identified by X-ray crystallography (Pavletich & Pabo 1991) and database analysis (Jacobs 1992).

25 As has been pointed out before (Berg 1992 Proc. Natl Acad. Sci. USA 89, 11109–11110), guanine has a particularly important role in zinc finger-DNA interactions. When present at the 5' (e.g. Table 1c-i) or 3' (e.g. Table 1m-o) end of a triplet, G selects fingers with Arg at position +6 or -1 30 of the α -helix respectively. When G is present in the middle position of a triplet (e.g. Table 1b), the preferred amino acid at position +3 is His. Occasionally, G at the 5' end of a triplet selects Ser or Thr at +6 (e.g. Table 1p). Since G can only be specified absolutely by Arg (Seeman et al., 1976 Proc. Nat. 35 Acad. Sci. USA 73, 804–808), this is the most common determinant at -1 and +6. One can expect this type of contact to be a bidentate hydrogen bonding interaction as seen in the crystal structures of Zif268 (Pavletich & Pabo 1991 Science 252, 809–817) and tramtrack (Fairall et al., 40 1993). In these structures, and in almost all of the selected fingers in which Arg recognises G at the 3' end, Asp occurs at position +2 to buttress the long Arg side chain (e.g. Table 1o,p). When position -1 is not Arg, Asp rarely occurs at +2, suggesting that in this case any other contacts it might make 45 with the second DNA strand do not contribute significantly to the stability of the protein-DNA complex.

Adenine is also an important determinant of sequence specificity, recognised almost exclusively by Asn or Gln which again are able to make bidentate contacts (Seeman et 50 al., 1976). When A is present at the 3' end of a triplet, Gln is often selected at position -1 of the α -helix, accompanied by small aliphatic residues at +2 (e.g. Table 1b). Adenine in the middle of the triplet strongly selects Asn at +3 (e.g. Table 1c-e), except in the triplet CAG (Table 1a) which selected 55 only two types of finger, both with His at +3 (one being the wild-type Zif268 which contaminated the library during this experiment). The triplets ACG (Table 1j) and ATG (Table 1k), which have A at the 5' end, also returned oligoclonal mixtures of phage, the majority of which were of one clone 60 with Asn at +6.

In theory, cytosine and thymine cannot reliably be discriminated by a hydrogen bonding amino acid side chain in the major groove (Seeman et al., 1976). Nevertheless, C in the 3' position of a triplet shows a marked preference for Asp 65 or Glu at position -1, together with Arg at +1 (e.g. Table 1e-g). Asp is also sometimes selected at +3 and +6 when C is in the middle (e.g. Table 1o) and 5' (e.g. Table 1a) position

respectively. Although Asp can accept a hydrogen bond from the amino group of C, one should note that the positive molecular charge of C in the major groove (Hunter 1993 *J. Mol. Biol.* 230, 1025–1054) will favour an interaction with Asp regardless of hydrogen bonding contacts. However, C in the middle position most frequently selects Thr (e.g. Table 1*i*), Val or Leu (e.g. Table 1*o*) at +3. Similarly, T in the middle position most often selects Ser (e.g. Table 1*i*), Ala or Val (e.g. Table 1*p*) at +3. The aliphatic amino acids are unable to make hydrogen bonds but Ala probably has a hydrophobic interaction with the methyl group of T, whereas a longer side chain such as Leu can exclude T and pack against the ring of C. When T is at the 5' end of triplet, Ser and Thr are selected at +6 (as is occasionally the case for G at the 5' end). Thymine at the 3' end of a triplet selects a variety of polar amino acids at -1 (e.g. Table 1*d*), and occasionally returns fingers with Ser at +2 (e.g. Table 1*a*) which could make a contact as seen in the traintrack crystal structure (Fairall et al., 1993).

Limitations of phage display. From Table 1 it can be seen that a consensus or bias usually occurs in two of the three primary positions (-1, +3 and +6) for any family of equivalent fingers, suggesting that in many cases phage selection is by virtue of only two base contacts per finger, as is observed in the Zif268 crystal structure (avletich & Pabo 1991). Accordingly, identical finger sequences are often returned by DNA sequences differing by one base in the central triplet. One reason for this is that the phage display selection, being essentially purification by affinity, can yield zinc fingers which bind equally tightly to a number of DNA triplets and so are unable to discriminate. Secondly, since complex formation is governed by the law of mass action, affinity selection can favour those clones whose representation in the library is greatest even though their true affinity for DNA is less than that of other clones less abundant in the library. Phage display selection by affinity is therefore of limited value in distinguishing between permissive and specific interactions beyond those base contacts necessary to stabilise the complex. Thus in the absence of competition from fingers which are able to bind specifically to a given DNA, the tightest non-specific complexes will be selected from the phage library. Consequently, results obtained by phage display selection from a library must be confirmed by specificity assays, particularly when that library is of limited size.

Conclusion. The amino acid sequence biases observed within a family of functionally equivalent zinc fingers indicate that, of the a-helical positions randomised in this study, only three primary (-1, +3 and +6) and one auxiliary (+2) positions are involved in recognition of DNA. Moreover, a limited set of amino acids are to be found at those positions, and it is presumed that these make contacts to bases. The indications therefore are that a code can be derived to describe zinc finger-DNA interactions. At this stage however, although sequence homologies are strongly suggestive of amino acid preferences for particular base-pairs, one cannot confidently deduce such rules until the specificity of individual fingers for DNA triplets is confirmed. The inventors therefore defer making a summary table of these preferences until the following example, in which is described how randomised DNA binding sites can be used to this end.

While this work was in progress, a paper by Rebar and Pabo was published (Rebar & Pabo 1994 *Science* 263, 671–673) in which phage display was also used to select zinc fingers with new DNA-binding specificities. These authors constructed a library in which the first finger of

Zif268 is randomised, and screened with tetranucleotides to take into account end effects such as additional contacts from variants of this finger. Only 4 positions (-1, +2, +3 and +6) were randomised, chosen on the basis of the earlier X-ray crystal structures. The results presented above, in which more positions were randomised, to some extent justifies Rebar and Pabo's use of the four random positions without apparent loss of effect, although further selections may reveal that the library is compromised. However, randomising only four positions decreases the theoretical library size so that full degeneracy can be achieved in practice. Nevertheless the inventors found that the results obtained by Rebar and Pabo by screening their complete library with two variant Zif268 operators, are in agreement with their conclusions derived from an incomplete library. On the one hand this again highlights the versatility of zinc fingers but, remarkably, so far both studies have been unable to produce fingers which bind to the sequence CCT. It will be interesting to see whether sequence biases such as we have detected would be revealed, if more selections were performed using Rebar and Pabo's library. In any case, it would be desirable to investigate the effects on selections of using different numbers of randomised positions in more complete libraries than have been used so far.

The original position or context of the randomised finger in the phage display library might bear on the efficacy of selected fingers when incorporated into a new DNA-binding domain. Selections from a library of the outer fingers of a three finger peptide (Rebar & Pabo, 1994 *Science* 263, 671–673; Jamieson et al., 1994 *Biochemistry* 33, 5689–5695) are capable of producing fingers which bind DNA in various different modes, while selections from a library of the middle finger should produce motifs which are more constrained. Accordingly, Rebar and Pabo do not assume that the first finger of Zif268 will always bind a triplet, and screened with a tetranucleotide binding site to allow for different binding modes. Thus motifs selected from libraries of the outer fingers might prove less amenable to the assembly of multifinger proteins, since binding of these fingers could be perturbed on constraining them to a particular binding mode, as would be the case for fingers which had to occupy the middle position of an assembled three-finger protein. In contrast, motifs selected from libraries of the middle finger, having been originally constrained, will presumably be able to preserve their mode of binding even when placed in the outer positions of an assembled DNA-binding domain.

FIGS. 10A–10C show different strategies for the design of tailored zinc finger proteins. (A) A three-finger DNA-binding motif is selected en bloc from a library of three randomised fingers. (B) A three-finger DNA-binding motif is assembled out of independently selected fingers from a library of one randomised finger (e.g. the middle finger of Zif268). (C) A three-finger DNA-binding motif is assembled out of independently selected fingers from three positionally specified libraries of randomised zinc fingers.

FIG. 11 illustrates the strategy of combinatorial assembly followed by en bloc selection. Groups of triplet-specific zinc fingers (A) isolated by phage display selection are assembled in random combinations and re-displayed on phage (B). A full-length target site (C) is used to select en bloc the most favourable combination of fingers (D).

EXAMPLE 2

This example describes a new technique to deal efficiently with the selection of a DNA binding site for a given zinc finger (essentially the converse of example 1). This is

desirable as a safeguard against spurious selections based on the screening of display libraries. This may be done by screening against libraries of DNA triplet binding sites randomised in two positions but having one base fixed in the third position. The technique is applied here to determine the specificity of fingers previously selected by phage display. The inventors found that some of these fingers are able to specify a unique base in each position of the cognate triplet. This is further illustrated by examples of fingers which can discriminate between closely related triplets as measured by their respective equilibrium dissociation constants. Comparing the amino acid sequences of fingers which specify a particular base in a triplet, we infer that in most instances, sequence specific binding of zinc fingers to DNA can be achieved using a small set of amino acid-base contacts amenable to a code.

One can determine the optimal binding sites of these (and other) proteins, by selection from libraries of randomised DNA. This approach, the principle of which is essentially the converse of zinc finger phage display, would provide an equally informative database from which the same rules can be independently deduced. However until now, the favoured method for binding site determination (involving iterative selection and amplification of target DNA followed by sequencing), has been a laborious process not conveniently applicable to the analysis of a large database (Thiesen & Bach 1990 Nucleic Acids Res. 18, 3203-3209; Pollock & Treisman 1990 Nucleic Acids Res. 18, 6197-6204).

This example presents a convenient and rapid new method which can reveal the optimal binding site(s) of a DNA binding protein by single step selection from small libraries and use this to check the binding site preferences of those zinc fingers selected previously by phage display. For this application, the inventors have used 12 different mini-libraries of the Zif268 binding site, each one with the central triplet having one position defined with a particular base pair and the other two positions randomised. Each library therefore comprises 16 oligonucleotides and offers a number of potential binding sites to the middle finger, provided that the latter can tolerate the defined base pair. Each zinc finger phage is screened against all 12 libraries individually immobilised in wells of a microtitre plate, and binding is detected by an enzyme immunoassay. Thus a pattern of acceptable bases at each position is disclosed, which the inventors term a "binding site signature". The information contained in a binding site signature encompasses the repertoire of binding sites recognised by a zinc finger.

The binding site signatures obtained, using zinc finger phage selected as described in example 1, reveal that the selection has yielded some highly sequence-specific zinc finger binding motifs which discriminate at all three positions of a triplet. From measurements of equilibrium dissociation constants it is found that these fingers bind tightly to the triplets indicated in their signatures, and discriminate against closely related sites (usually by at least a factor of ten). The binding site signatures allow progress towards a specificity code for the interactions of zinc fingers with DNA.

MATERIALS AND METHODS

Binding site signatures. Flexible flat-bottomed 96-well microtitre plates (Falcon) were coated overnight at 4° C. with streptavidin (0.1 mg/ml in 0.1M NaHCO₃, pH 8.6, 0.3% NaN₃). Wells were blocked for one hour with PBS/Zn (PBS, 50 μM Zn (CH₃COO)₂) containing 2% fat-free dried milk (Marvel), washed 3 times with PBS/Zn containing 0.1% TWEEN, and another 3 times with PBS/Zn. The "bound" strand of each oligonucleotide library was made syntheti-

cally and the other strand extended from a 5'-biotinylated universal primer using DNA polymerase I (Klenow fragment). Fill-in reactions were added to wells (0.8 pmole DNA library in each) in PBS/Zn for 15 minutes, then washed once with PBS/Zn containing 0.1% TWEEN, and once again with PBS/Zn. Overnight bacterial cultures each containing a selected zinc finger phage were grown in 2xTY containing 50 mM Zn(CH₃COO)₂ and 15 μg/ml tetracycline at 30° C. Culture supernatants containing phage were diluted tenfold by the addition of PBS/Zn containing 2% fat-free dried milk (Marvel), 1% TWEEN and 20 μg/ml sonicated salmon sperm DNA. Diluted phage solutions (50 μl) were applied to wells and binding, allowed to proceed for one hour at 20° C. Unbound phage were removed by washing 5 times with PBS/Zn containing 1% TWEEN, and then 3 times with PBS/Zn. Bound phage were detected as described previously (Griffiths et al., 1994 EMBO J. In press), or using HRP-conjugated anti-M13 IgG (Pharmacia), and quantitated using SOFTMAX 2.32 (Molecular Devices Corp).

The results are shown in FIG. 4, which gives the binding site signatures of individual zinc finger phage. The figure represents binding of zinc finger phage to randomised DNA immobilised in the wells of microtitre plates. To test each zinc finger phage against each oligonucleotide library (see above), DNA libraries are applied to columns of wells (down the plate), while rows of wells (across the plate) contain equal volumes of a solution of a zinc finger phage. The identity of each library is given as the middle triplet of the "bound" strand of Zif268 operator, where N represents a mixture of all 4 nucleotides. The zinc finger phage is specified by the sequence of the variable region of the middle finger, numbered relative to the first helical residue (position 1), and the three primary recognition positions are highlighted. Bound phage are detected by an enzyme immunoassay. The approximate strength of binding is indicated by a grey scale proportional to the enzyme activity. From the pattern of binding to DNA libraries, called the "signature" of each clone, one or a small number of binding sites can be read off and these are written on the right of the figure.

Determination of apparent equilibrium dissociation constants. Overnight bacterial cultures were grown in 2xTY/Zn/Tet at 30° C. Culture supernatants containing phage were diluted twofold by the addition of PBS/Zn containing 4% fat-free dried milk (Marvel), 2% Tween and 40 μg/ml sonicated salmon sperm DNA. Binding reactions, containing appropriate concentrations of specific 5'-biotinylated DNA and equal volumes of zinc finger phage solution, were allowed to equilibrate for 1 h at 20° C. All DNA was captured on streptavidin-coated paramagnetic beads (500 μg per well) which were subsequently washed 6 times with PBS/Zn containing 1% Tween and then 3 times with PBS/Zn. Bound phage were detected using HRP-conjugated anti-M13 IgG (Pharmacia) and developed as described (Griffiths et al., 1994). Optical densities were quantitated using SOFTMAX 2.32 (Molecular Devices Corp).

The results are shown in FIGS. 5A-5F, which are series of graphs of fractional saturation against concentration of DNA (nM). The two outer fingers carry the native sequence, as do the two cognate outer DNA triplets. The sequence of amino acids occupying helical positions -1 to +9 of the varied finger are shown in each case. The graphs show that the middle finger can discriminate closely related triplets, usually by a factor of ten. The graphs allowed the determination of apparent equilibrium dissociation constants, as below.

Estimations of the K_d are by fitting to the equation K_d=[DNA].[P]/[DNA.P], using the software package

KALEIDAGRAPH™ Version 2.0 programme (Abelbeck Software). Owing to the sensitivity of the ELISA used to detect protein-DNA complex, the inventors were able to use zinc finger phage concentrations far below those of the DNA, as is required for accurate calculations of the K_d . The technique used here has the advantage that while the concentration of DNA (variable) must be known accurately, that of the zinc fingers (constant) need not be known (Choo & Kiug 1993 Nucleic Acids Res. 21, 3341–3346). This circumvents the problem of calculating the number of zinc finger peptides expressed on the tip of each phage, although since only 10–20% of the gene III protein (pIII) carries such peptides one would expect on average less than one copy per phage. Binding is performed in solution to prevent any effects caused by the avidity (Marks et al., 1992) of phage for DNA immobilised on a surface. Moreover, in this case measurements of K_d by ELISA are made possible since equilibrium is reached in solution prior to capture on the solid phase.

RESULTS AND DISCUSSION

The binding site signature of the second finger of Zif268. The top row of FIG. 4 shows the signature of the second finger of wild type Zif268. From the pattern of strong signals indicating binding to oligonucleotide libraries having GNN, TNN, NGN and NNG as the middle triplet, it emerges that the optimal binding site for this finger is T/G,G,G, in accord with the published consensus sequence (Christy & Nathans 1989 Proc. Natl. Acad. Sc. USA 86, 8737–8741). This has implications for the interpretation of the X-ray crystal structure of Zif268 solved in complex with consensus operator having TGG as the middle triplet (Pavletich & Pabo 1991). For instance, His at position +3 of the middle finger was modelled as donating a hydrogen bond to N7 of G, suggesting an equivalent contact to be possible with N7 of A, but from the binding site signature we can see that there is discrimination against A. This implies that the His may prefer to make a hydrogen bond to O6 of G or a bifurcated hydrogen bond to both O6 and N7, or that a steric clash with the amino group of A may prevent a tight-interaction with this base. Thus by considering the stereochemistry of double helical DNA, binding site signatures can give insight into the details of zinc finger-DNA interactions.

Amino acid-base contacts in zinc finger-DNA complexes deduced from binding site signatures. The binding site signatures of other zinc fingers reveal that the phage selections performed in example 1 yielded highly sequence-specific DNA binding proteins. Some of these are able to specify a unique sequence for the middle triplet of a variant Zif268 binding site, and are therefore more specific than is Zif268 itself for its consensus site. Moreover, one can identify the fingers which recognise a particular oligonucleotide library, that is to say a specific base at a defined position, by looking down the columns of FIG. 4. By comparing the amino acid sequences of these fingers one can identify any residues which have genuine preferences for particular bases on bound DNA. With a few exceptions, these are as previously predicted on the basis of phage display, and are summarised in Table 2.

Table 2 summarises frequently observed amino acid-base contacts in interactions of selected zinc fingers with DNA. The given contacts comprise a "syllabic" recognition code for appropriate triplets. Cognate amino acids and their positions in the α -helix are entered in a matrix relating each base to each position of a triplet. Auxiliary amino acids from position +2 can enhance or modulate specificity of amino acids at position -1 and these are listed as pairs. Ser or Thr at position +6 permit Asp +2 of the following finger (denoted

Asp ++2) to specify both G and T indirectly, and the pairs are listed. The specificity of Ser+3 for T and Thr+3 for C may be interchangeable in rare instances while Val+3 appears to be consistently ambiguous.

TABLE 2

POSITION IN TRIPLET				
	5'	MIDDLE	3'	
10	G	Arg +6 Ser +6/Asp ++2 Thr +6/Asp ++2	His +3	Arg -1/Asp +2
15	A		Asn +3	Gln -1/Ala +2
	T	Ser +6/Asp ++2 Thr +6/Asp ++2	Ala +3 Ser +3 Val +3	Asn -1 Gln -1/Ser +2
	C		Asp +3 Leu +3 Thr +3 Val +3	Asp -1

The binding site signatures also reveal an important feature of the phage display library which is important to the interpretation of the selection results. All the fingers in our panel, regardless of the amino acid present at position +6, are able to recognise G or both G and T at the 5' end of a triplet. The probable explanation for this is that the 5' position of the middle triplet is fixed as either G or T by a contact from the invariant Asp at position +2 of finger 3 to the partner of either base on the complementary strand, analogous to those seen in the Zif268 (Pavletich & Pabo 1991 Science 252, 809–817) and tramtrack (Fairall et al., 1993) crystal structures (a contact to NH₂ of C or A respectively in the major groove). Therefore Asp at position +2 of finger 3 is dominant over the amino acid present at position +6 of the middle finger, precluding the possibility of recognition of A or C at the 5' position. Future libraries must be designed with this interaction omitted or the position varied. Interestingly, given the framework of the conserved regions of the three fingers, one can identify a rule in the second finger which specifies a frequent interaction with both G and T, viz the occurrence of Ser or Thr at position +6, which may donate a hydrogen bond to either base.

Modulation of base recognition by auxiliary positions. As noted above, position +2 is able to specify the base directly 3' of the 'cognate triplet', and can thus work in conjunction with position +6 of the preceding finger. The binding site signatures, whilst pointing to amino acid-base contacts from the three primary positions, indicate that auxiliary positions can play other parts in base recognition. A clear case in point is Gln at position -1, which is specific for A at the 3' end of a triplet when position +2 is a small non-polar amino acid such as Ala, though specific for T when polar residues such as Ser are at position +2. The strong correlation between Arg at position -1 and Asp at position +2, the basis of which is understood from the X-ray crystal structures of zinc fingers, is another instance of interplay between these two positions. Thus the amino acid at position +2 is able to modulate or enhance the specificity of the amino acid at other positions.

At position +3, a different type of modulation is seen in the case of Thr and Val which most often prefer C in the middle position of a triplet, but in some zinc fingers are able to recognise both C and T. This ambiguity occurs possibly as a result of different hydrophobic interactions involving the methyl groups of these residues, and here a flexibility in the inclination of the finger rather than an effect from another position per se may be the cause of ambiguous reading.

Quantitative measurements of dissociation constants. The binding site signature of a zinc finger reveals its differential base preferences at a given concentration of DNA. As the concentration of DNA is altered, one can expect the binding site signature of any clone to change, being more distinctive at low [DNA], and becoming less so at higher [DNA] as the K_d of less favourable sites is approached and further bases become acceptable at each position of the triplet. Furthermore, because two base positions are randomly occupied in any one library of oligonucleotides, binding site signatures are not formally able to exclude the possibility of context dependence for some interactions. Therefore to supplement binding site signatures, which are essentially comparative, quantitative determinations of the equilibrium dissociation constant of each phage for different DNA binding sites are required. After phage display selection and binding site signatures, these are the third and definitive stage in assessing the specificity of zinc fingers.

Examples of such studies presented in FIGS. 5A-5F, which reveal that zinc finger phages bind the operators indicated in their binding site signatures with K_d s in the range of 10^{-8} - 10^{-9} M, and can discriminate against closely related binding sites by factors greater than an order of magnitude. Indeed FIG. 5 shows such differences in affinity for binding sites which differ in only one out of nine base pairs. Since the zinc fingers in our panel were selected from a library by non-competitive affinity purification, there is the possibility that fingers which are even more discriminatory can be isolated using a competitive selection process.

Measurements of dissociation constants allow different triplets to be ranked in order of preference according to the strength of binding. The examples here indicate that the contacts from either position -1 or +3) can contribute to discrimination. Also, the ambiguity in certain binding site signatures referred to above can be shown to have a basis in the equal affinity of certain figures for closely related triplets. This is demonstrated by the K_d s of the finger containing the amino acid sequence RGDALTSHER (Seq ID No. 100) for the triple TTG and GTG.

A code for zinc finger-DNA recognition. One would expect that the versatility of the zinc finger motif will have allowed evolution to develop various modes of binding to DNA (and even to RNA), which will be too diverse to fall under the scope of a single code. However, although a code may not apply to all zinc finger-DNA interactions, there is now convincing evidence that a code applies to a substantial subset. This code will fall short of being able to predict unfailingly the DNA binding site preference of any given zinc finger from its amino acid sequence, but may yet be sufficiently comprehensive to allow the design of zinc fingers with specificity for a given DNA sequence.

Using the selection methods of phage display (as described above) and of binding site signatures it is found that in the case of Zif268-like zinc fingers, DNA recognition involves four fixed principal (three primary and one auxiliary) positions on the α -helix, from where a limited and specific set of amino acid-base contacts result in recognition of a variety of DNA triplets. In other words, a code can describe the interactions of zinc fingers with DNA. Towards this code, one can propose amino acid-base contacts for almost all the entries in a matrix relating each base to each position of a triplet (Table 2). Where there is overlap, the results presented here complement those of Desjarlais and Berg who have derived similar rules by altering zinc finger specificity using database-guided mutagenesis (Desjarlais & Berg 1992 Proc Natl. Acad. Sci. USA 89, 7345-7349; Desjarlais & Berg 1993 Proc. Natl. Acad. Sci. USA 90, 2256-2260).

Combinatorial use of the coded contacts. The individual base contacts listed in Table 2, though part of a code, may not always result in sequence specific binding to the expected base triplet when used in any combination. In the first instance one must be aware of the possibility that zinc fingers may not be able to recognise certain combinations of bases in some triplets by use of this code, or even at all. Otherwise, the majority of inconsistencies may be accounted for by considering variations in the inclination of the trident reading head of a zinc finger with respect to the triplet with which it is interacting. It appears that the identity of an amino acid at any one α -helical position is attuned to the identity of the residues at the other two positions to allow three base contacts to occur simultaneously. Therefore, for example, in order that Ala may pick out T in the triplet GTG, Arg must not be used to recognise G from position +6, since this would distance the former too far from the DNA (see for example the finger containing the amino acid sequence RGDALTSHER). (Seq ID No. 100). Secondly, since the pitch of the α -helix is 3.6 amino acids per turn, positions -1, +3 and +6 are not an integral number of turns apart, so that position +3 is nearer to the DNA than are -1 or +6. Hence, for example, short amino acids such as His and Asn, rather than the longer Arg and Gln, are used for the recognition of purines in the middle position of a triplet.

As a consequence of these distance effects one might say that the code is not really "alphabetic" (always identical amino acid:base contact) but rather "syllabic" (use of a small repertoire of amino acid:base contacts). An alphabetic code would involve only four rules, but syllabicity adds an additional level of complexity, since systematic combinations of rules comprise the code. Nevertheless, the recognition of each triplet is still best described by a code of syllables, rather than a catalogue of "logograms" (idiosyncratic amino acid:base contact depending on triplet).

Conclusions. The "syllabic" code of interactions with DNA is made possible by the versatile framework of the zinc finger: this allows an adaptability at the interface with DNA by slight changes of orientation, which in turn maintains a stoichiometry of one coplanar amino acid per base-pair in many different complexes. Given this mode of interaction between amino acids and bases it is to be expected that recognition of G and A by Arg and Asn/Gln respectively are important features of the code; but remarkably other interactions can be more discriminatory than was anticipated (Seeman et al., 1976). Conversely, it is clear that degeneracy can be programmed in the zinc fingers in varying degrees allowing for intricate interactions with different regulatory DNA sequences (Harrison & Travers, 1990; Christy & Nathans, 1989). One can see how this principle makes possible the regulation of differential gene expression by a limited set of transcription factors.

As already noted above, the versatility of the finger motif will likely allow other modes of binding to DNA. Similarly, one must take into account the malleability of nucleic acids such as is observed in Fairall et al., where a deformation of the double helix at a flexible base step allows a direct contact from Ser at position +2 of finger 1 to a T at the 3' position of the cognate triplet. Even in our selections there are instances of fingers whose binding mode is obscure, and may require structural analyses for clarification. Thus, water may be seen to play an important role, for example where short side chains such as Asp, Asn or Ser interact with bases from position -1 (Qian et al., 1993 J. Am. Chem. Soc. 115, 1189-1190; Shakked et al., 1994 Nature (London) 368, 469-478).

Eventually, it might be possible to develop a number of codes describing zinc finger binding to DNA, which could

predict the binding site preferences of some zinc fingers from their amino acid sequence. The functional amino acids selected at positions -1, +3 and to an extent +6 in this study, are very frequently observed at the same positions in naturally occurring fingers (e.g. see FIG. 4. of Desjarlais and Berg 1992 Proteins 12, 101-104) supporting the existence of coded contacts from these three positions. However, the lack of definitive predictive methods is not a serious practical limitation as current laboratory techniques (here and in Thiesen & Bach 1990 and Pollock & Treisman 1990) will allow the identification of binding sites for a given DNA-binding protein. Rather, one can apply phage selection and a knowledge of the recognition rules to the converse problem, namely the design of proteins to bind predetermined DNA sites.

Prospects for the design of DNA-binding proteins. The ability to manipulate the sequence specificity of zinc fingers implies that we are on the eve of designing DNA-binding proteins with desired specificity for applications in medicine and research (Desjarlais & Berg, 1993; Rebar & Pabo, 1994). This is possible because, by contrast to all other DNA-binding motifs, we can avail ourselves of the modular nature of the zinc finger, since DNA sites can be recognised by appropriate combinations of independently acting fingers linked in tandem.

The coded interactions of zinc fingers with DNA can be used to model the specificity of individual zinc fingers de novo, or more likely in conjunction with phage display selection of suitable candidates. In this way, according to requirements, one could modulate the affinity for a given binding site, or even engineer an appropriate degree of indiscrimination at particular base positions. Moreover, the additive effect of multiply repeated domains offers the opportunity to bind specifically and tightly to extended, and hence very rare, genomic loci. Thus zinc finger proteins might well be a good alternative to the use of antisense nucleic acids in suppressing or modifying the action of a given gene, whether normal or mutant. To this end, extra functions could be introduced to these DNA binding domains by appending suitable natural or synthetic effectors.

EXAMPLE 3

From the evidence presented in the preceding examples, the inventors propose that specific DNA-binding proteins comprising zinc fingers can be "made to measure". To demonstrate their potential the inventors have created a three finger polypeptide able to bind site-specifically to a unique 9 bp region of a BCR-ABL fusion oncogene and to discriminate it from the parent genomic sequences (Kurzrock et al., 1988 N. Engl. J. Med. 319, 990-998). Using transformed cells in culture as a model, it is shown that binding to the target oncogene in chromosomal DNA is possible, resulting in blockage of transcription. Consequently, murine cells made growth factor-independent by the action of the oncogene (Daley et al., 1988 Proc. Natl. Acad. Sci. U.S.A. 85, 9312-9316) are found to revert to factor dependence on transient transfection with a vector expressing the designed zinc finger polypeptide.

DNA-binding proteins designed to recognise specific DNA sequences could be incorporated in chimeric transcription factors, recombinases, nucleases etc. for a wide range of applications. The inventors have shown that zinc finger mini-domains can discriminate between closely related DNA triplets, and have proposed that they can be linked together to form domains for the specific recognition of longer DNA sequences. One interesting possibility for the use of such protein domains is to target selectively genetic

differences in pathogens or transformed cells. Here one such application is described.

There exist a set of human leukaemias in which a reciprocal chromosomal translocation t(9;22) (q34;q11) result in a truncated chromosome 22, the Philadelphia chromosome (Ph1), encoding at the breakpoint a fusion of sequences from the c-ABL protooncogene (Bartram et al., 1983 Nature 306, 277-280) and the BCR gene (Groffen et al., 1984 Cell 36, 93-99). In chronic myelogenous leukaemia (CML), the breakpoints usually occur in the first intron of the c-ABL gene and in the breakpoint cluster region of the BCR gene (Shtivelman et al., 1985 Nature 315, 550-554), and give rise to a p210^{BCR-ABL} gene product (Konopka et al., 1984 Cell 37, 1035-1042). Alternatively, in acute lymphoblastic leukaemia (ALL), the breakpoints usually occur in the first introns of both BCR and c-ABL (Hermans et al., 1987 Cell 51, 33-40), and result in a p190^{BCR-ABL} gene product (FIG. 6) (Kurzrock et al., 1987 Nature 325, 631-635).

FIG. 6 shows the nucleotide sequences (Seq ID No.s 9-11) of the fusion point between BCR and ABL sequences in p190 cDNA, and of the corresponding exon boundaries in the BCR and c-ABL genes. Exon sequences are written in capital letters while introns are given in lowercase. Line 1 shows p190^{BCR-ABL} cDNA; line 2 the BCR genomic sequence at junction of exon 1 and intron 1; and line 3 the ABL genomic sequence at junction of intron 1 and exon 2 (Hermans et al. 1987). The 9 bp sequence in the p190^{BCR-ABL} cDNA used as a target is underlined, as are the homologous sequences in genomic BCR and c-ABL.

Facsimiles of these rearranged genes act as dominant transforming oncogenes in cell culture (Daley et al., 1988) and transgenic mice (Heisterkamp et al., 1990 Nature 344, 251-253). Like their genomic counterparts, the cDNAs bear a unique nucleotide sequence at the fusion point of the BCR and c-ABL genes, which can be recognised at the DNA level by a site-specific DNA-binding protein. The present inventors have designed such a protein to recognise the unique fusion site in the p190^{BCR-ABL} cDNA. This fusion is obviously distinct from the breakpoints in the spontaneous genomic translocations, which are thought to be variable among patients. Although the design of such peptides has implications for cancer research, the primary aim here is to prove the principle of protein design, and to assess the feasibility of *in vivo* binding to chromosomal DNA in available model systems.

A nine base-pair target sequence (GCA, GAA, GCC) for a three zinc finger peptide was chosen which spanned the fusion point of the p190^{BCR-ABL} cDNA (Hermans et al., 1987). The three triplets forming this binding site were each used to screen a zinc finger phage library over three rounds as described above in example 1. The selected fingers were then analysed by binding site signatures to reveal their preferred triplet, and mutations to improve specificity were made to the finger selected for binding to GCA. A phage display mini-library of putative BCR-ABL-binding three-finger proteins was cloned in fd phage, comprising six possible combinations of the six selected or designed fingers (1A, 1B; 2A; 3A, 3B and 3C) linked in the appropriate order. These fingers are illustrated in FIG. 7 (Seq ID No.s 12-17).

In FIG. 7 regions of secondary structure are underlined below the list, while residue positions are given above, relative to the first position of the α -helix (position 1). Zinc finger phages were selected from a library of 2.6×10^6 variants, using three DNA binding sites each containing one of the triplets GCC, GAA or GCA. Binding site signatures (example 2) indicate that fingers 1A and 1B specify the triplet GCC, finger 2A specifies GAA, while the fingers

selected using the triplet GCA all prefer binding to GCT. Amongst the latter is finger 3A, the specificity of which we believed, on the basis of recognition rules, could be changed by a point mutation. Finger 3B, based on the selected finger 3A, but in which Gln at helical position +2 was altered to Ala should be specific for GCA. Finger 3C is an alternative version of finger 3A, in which the recognition of C is mediated by Asp+3 rather than by Thr+3.

The mini library was screened once with an oligonucleotide containing the 9 base-pair BCR-ABL target sequence to select for tight binding clones over weak binders and background vector phage. Because the library was small, the inventors did not include competitor DNA sequences for homologous regions of the genomic BCR and c-ABL genes but instead checked the selected clones for their ability to discriminate. It was found that although all the selected clones were able to bind the BCR-ABL target sequence and to discriminate between this and the genomic-BCR sequence, only a subset could discriminate against the c-ABL sequence which, at the junction between intron 1 and exon 2, has an 8/9 base-pair homology to the BCR-ABL target sequence (Hermans et al., 1987). Sequencing of the discriminating clones revealed two types of selected peptide, one with the composition 1A-2A-3B and the other with 1B-2A-3B. Thus both peptides carried the third finger (3B) which was specifically designed against the triplet GCA but peptide 1A-2A-3B was able to bind to the BCR-ABL target sequence with higher affinity than was peptide 1B-2A-3B.

The peptide 1A-2A-3B, henceforth referred to as the anti-BCR-ABL peptide, was used in further experiments. The anti-BCR-ABL peptide has an apparent equilibrium dissociation constant (K_d) of $6.2+/-0.4\times10^{-7}$ M for the p190^{BCR-ABL} cDNA sequence in vitro, and discriminates against the similar sequences found in genomic BCR and c-ABL DNA, by factors greater than an order of magnitude (FIG. 8). Referring to FIG. 8, (which illustrates discrimination in the binding of the anti-BCR-ABL peptide to its p190^{BCR-ABL} target site and to like regions of genomic BCR and c-ABL), the graph shows binding (measured as an $A_{450-650}$) at various [DNA]. Binding reactions and complex detection by enzyme immunoassay were performed as described previously, and a full curve analysis was used in calculations of the K_d (Choo & Kiug 1993). The DNA used were oligonucleotides spanning 9bp either side of the fusion point in the cDNA or the exon boundaries. The anti-BCR-ABL peptide binds to its intended target site with a $K_d=6.2+/-0.4\times10^{-7}$ M, and is able to discriminate against genomic BCR and c-ABL sequences, though the latter differs by only one base pair in the bound 9 bp region. The measured dissociation constant is higher than that of three-finger peptides from naturally occurring proteins such as Spl (Kadonga et al., 1987 Cell 51, 1079-1090) or Zif268 (Christy et al., 1988), which have K_d s in the range of 10^{-9} M, but rather is comparable to that of the two fingers from the tramtrack (ttk) protein (Fairall et al., 1992). However, the affinity of the anti-BCR-ABL peptide could be refined, if desired, by site-directed mutations or by "affinity maturation" of a phage display library (Hawkins et al., 1992 J. Mol. Biol. 226, 889-896).

Having established DNA discrimination in vitro, the inventors wished to test whether the anti-BCR-ABL peptide was capable of site-specific DNA-binding in vivo. The peptide was fused to the VP16 activation domain from herpes simplex virus (Fields 1993 Methods 5, 116-124) and used in transient transfection assays (FIG. 9) to drive production of a CAT (chloramphenicol acetyl transferase) reporter gene from a binding site upstream of the TATA box

(Gorman et al., Mol. Cell. Biol. 2, 1044-1051). In detail the experiment was performed thus: reporter plasmids pMCAT6BA, pMCAT6A, and pMCAT6B, were constructed by inserting 6 copies of the p190^{BCR-ABL} target site (CGCAGAAAGCC, SEQ ID No. 121), the c-ABL second exon-intron junction sequence (TCCAGAAAGCC, Seq ID No. 122), or the BCR first exon-intron junction sequence (CGCAGGTGAG Seq ID No. 123) respectively, into pMCAT3 (Luscher et al., 1989 Genes Dev. 3, 1507-1517). The anti-BCR-ABL/VP16 expression vector was generated by inserting the in-frame fusion between the activation domain of herpes simplex virus VP16 (Fields 1993) and the Zn finger peptide in the pEF-BOS vector (Mizushima & Shigezaku 1990 Nucl. Acids Res. 18, 5322). C3H10T1/2 cells were transiently co-transfected with 10 mg of reporter plasmid and 10 mg of expression vector. RSVL (de Wet et al., 1987 Mol. Cell Biol. 7, 725-737), which contains the Rous sarcoma virus long terminal repeat linked to luciferase, was used as an internal control to normalise for differences in transfection efficiency. Cells were transfected by the calcium phosphate precipitation method and CAT assays performed as described (Sanchez-Garcia et al., 1993 EMBO J. 12, 4243-4250). Plasmid pGSEC, which has five consensus 17-mer GAL4-binding sites upstream from the minimal promoter of the adenovirus Elb TATA box, and pMIVP16 vector, which encodes an in-frame fusion between the DNA-binding domain of GAL4 and the activation domain of herpes simplex virus VP16, were used as a positive control (Sadowski et al., 1992 Gene 118, 137-141).

C3H10T1/2 cells were transiently cotransfected with a CAT reporter plasmid and an anti-BCR-ABL/VP16 expression vector (pZNIA).

A specific (thirty-fold) increase in CAT activity was observed in cells cotransfected with reporter plasmid bearing copies of the p190^{BCR-ABL} cDNA target site, compared to a barely detectable increase in cells cotransfected with reporter plasmid bearing copies of either the BCR or c-ABL semihomologous sequences, indicating in vivo binding.

The selective stimulation of transcription indicates convincingly that highly site-specific DNA-binding can occur in vivo. However, while transient transfections assay binding to plasmid DNA, the true target site for this and most other DNA-binding proteins is in genomic DNA. This might well present significant problems, not least since this DNA is physically separated from the cytosol by the nuclear membrane, but also since it may be packaged within chromatin.

To study whether genomic targeting is possible, a construct was made in which the anti-BCR-ABL peptide was flanked at the N-terminus with the nuclear localisation signal from the large T antigen of SV40 virus (Kalderon et al., 1984 Cell 499-509), and at the C-terminus with an 11 amino acid c-myc epitope tag recognisable by the 9E10 antibody (Evan et al., 1985 Mol. Cell. Biol. 5, 3610-3616). This construct was used to transiently transfet the IL-3-dependent murine cell line Ba/F3 (Palacios & Steinmetz 1985 Cell 41, 727-734), or alternatively Ba/F3 +p190 and Ba/F3+p210 cell lines previously made IL-3-independent by integrated plasmid constructs expressing either p190^{BCR-ABL} or p210^{BCR-ABL}, respectively. Staining of the cells with the 9E10 antibody followed by a secondary fluorescent conjugate showed efficient nuclear localisation in those cells transfected with the anti-BCR-ABL peptide.

The experimental details were as follows: the anti-BCR-ABL expression vector was generated in the pEF-BOS vector (Mizushima & Shigezaku 1990), including an 11

amino acid c-myc epitope tag (EQKLISEEDLN, Seq ID No. 124) at the carboxy-terminal end, recognizable by the 9E10 antibody (Evan et al., 1985) and the nuclear localization signal (PKKKRKV, Seq ID No. 125) of the large T antigen of SV40 virus (Kalderon et al., 1984) at the amino-terminal end. Three glycine residues were introduced downstream of the nuclear localization signal as a spacer, to ensure exposure of the nuclear leader from the folded molecule. Ba/F3 cells were transfected with 25 mg of the anti-BCR-ABL expression construct tagged with the 9E10 c-myc epitope as described (Sanchez-Garcia & Rabbitts 1994 Proc. Natl. Acad. Sci. U.S.A. in press) and protein production analyzed 48 h later by immunofluorescence-labelling as follows. Cells were fixed in 4% (w/v) paraformaldehyde for 15 min, washed in phosphate-buffered saline (PBS), and permeabilized in methanol for 2 min. After blocking in 10% fetal calf serum in PBS for 30 min, the mouse 9E10 antibody was added. After a 30 min incubation at room temperature a fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (SIGMA) was added and incubated for a further 30 min. Fluorescent cells were visualized using a confocal scanning microscope (magnification, 200 \times). Immunofluorescence of Ba/F3+p190 and Ba/F3+p210 cells transiently transfected with the anti-bcr-abl expression vector and stained with the 9E10 antibody was done. Expression and nuclear localisation of the anti-BCR-ABL peptide was observed. In addition, transfected Ba/F3+p190 cells show chromatin condensation and nuclear fragmentation into small apoptotic bodies, but not either untransfected Ba/F3+p190 cells or transfected Ba/F3+p210 cells.

The efficiency of transient transfection, measured as the proportion of immunofluorescent cells in the population, was 15–20%. When IL-3 is withdrawn from tissue culture, a corresponding proportion of Ba/F3+p190 cells are found to have reverted to factor dependence and die, while Ba/F3+p210 cells are unaffected. The experimental details were as follows: cell lines Ba/F3, Ba/F3 +p190 and Ba/F3 +p210 were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. In the case of Ba/F3 cell line 10% WEHI-3B-conditioned medium was included as a source of IL-3. After the transfection with the anti-BCR-ABL expression vector, cells (5×10^5 /ml) were washed twice in serum-free medium and cultured in DMEM medium with 10% fetal bovine serum without WEHI-3B-conditioned medium. Percentage viability was determined by trypan blue exclusion. Data are expressed as means of triplicate cultures. The results are shown in graphical form in FIG. 9.

Immunofluorescence microscopy of transfected Ba/F3+p190 cells in the absence of IL-3 shows chromatin condensation and nuclear fragmentation into small apoptotic bodies, while the nuclei of Ba/F3+p210 cells remain intact. Northern blots of total cytoplasmic RNA from Ba/F3 +p190 cells transiently transfected with the anti-BCR-ABL peptide revealed reduced levels of p190^{BCR-ABL} mRNA relative to untransfected cells. By contrast, similarly transfected Ba/F3 +p210 cells showed no decrease in the levels of p210^{BCR-ABL} mRNA (FIG. 12). The blots were performed as follows: 10 mg of total cytoplasmic RNA, from the cells indicated, was glyoxylated and fractionated in 1.4% agarose gels in 10 mM NaPO₄ buffer, pH 7.0. After electrophoresis the gel was blotted onto HYBOND-N (Amersham), UV-cross linked and hybridized to an ³²P-labelled c-ABL probe. Autoradiography was for 14 h at -70° C. Loading was monitored by reprobing the filters with a mouse b-actin cDNA. Northern filter hybridisation analysis of Ba/F3+p190 and Ba/F3+p210 cell lines transfected with the anti-BCR-ABL expression

vector was done. When transfected with the anti-BCR-ABL expression vector, a specific downregulation of p190^{BCR-ABL} mRNA was seen in Ba/F3+p190 cells, while expression of p210^{BCR-ABL} was unaffected in Ba/F3+p210 cells.

In summary, the inventors have demonstrated that a DNA-binding protein designed to recognise a specific DNA sequence in vitro, is active in vivo where, directed to the nucleus by an appended localisation signal, it can bind its target sequence in chromosomal DNA. This is found on otherwise actively transcribing DNA, so presumably binding of the peptide blocks the path of the polymerase, causing stalling or abortion. The use of a specific polypeptide in this case to target intragenic sequences is reminiscent of antisense oligonucleotide- or ribozyme-based approaches to inhibiting the expression of selected genes (Stein & Cheng 1993 Science 261, 1004–1012). Like antisense oligonucleotides, zinc finger DNA-binding proteins can be tailored against genes altered by chromosomal translocations, or point mutations, as well as to regulatory sequences within genes. Also, like oligonucleotides which can be designed to repress transcription by triple helix formation in homopurine-homopyrimidine promoters (Cooney et al., 1988 Science 245, 725–730) DNA-binding proteins can bind to various unique regions outside genes, but in contrast they can direct gene expression by both up- or down-regulating, the initiation of transcription when fused to activation (Seipel et al., 1992 EMBO J. 11, 4961–4968) or repression domains (Herschbach et al., 1994 Nature 370, 309–311). In any case, by acting directly on any DNA, and by allowing fusion to a variety of protein effectors, tailored site-specific DNA-binding proteins have the potential to control gene expression, and indeed to manipulate the genetic material itself, in medicine and research.

EXAMPLE 4

The phage display zinc finger library described in the preceding examples could be considered sub-optimal in a number of ways:

- i) the library was much smaller than the theoretical maximum size;
- ii) the flanking fingers both recognised GCG triplets (in certain cases creating nearly symmetrical binding sites for the three zinc fingers, which enables the peptide to bind to the 'bottom' strand of DNA, thus evading the register of interactions we wished to set);
- iii) Asp+2 of finger three ("Asp++2") was dominant over the interactions of finger two (position +6) with the 5' base of the middle triplet;
- iv) not all amino acids were represented in the randomised positions.

In order to overcome these problems a new three-finger library was created in which:

- a) the middle finger is fully randomised in only four positions (-1, +3 and +6) so that the library size is smaller and all codons are represented. The library was cloned in the pCANTAB5E phagemid vector from Pharmacia, which allows higher transformation frequencies than the phage.
- b) the first and third fingers recognise the triplets GAC and GCA, respectively, making for a highly asymmetric binding site. Recognition of the 3'A in the latter triplet by finger three is mediated by Gln-1/Ala+2, the significance of which is that the short Ala+2 should not make contacts to the DNA (in particular with the 5' base of the middle triplet), thus alleviating the problem noted at (iii) above.

31

EXAMPLE 5

The human ras gene is susceptible to a number of different mutations, which can convert it into an oncogene. A ras oncogene is found in a large number of human cancers. One particular mutation is known as the G12V mutation (i.e. the polypeptide encoded by the mutant gene contains a substitution from glycine to valine). Because ras oncogenes are so common in human cancers, they are extremely significant targets for potential therapeutic methods.

A three finger protein has been designed which can recognise the G12V mutant of ras. The protein was produced using rational design based on the known specificity rules. In outline, a zinc finger framework (from one of the fingers selected to bind GCC) was modified by point mutations in position +3 to yield fingers recognising two additional different triplets. The finger recognising GCC and the two derivatives were cloned in pCANTAB5E and expressed on the surface of phage.

Originally, the G12V-binding peptide "r-BP" was to be selected from a small library of related proteins. The reason a library was to be used is that while it was clear to us what 8/9 of the amino acid:base contacts should be, it was not clear whether the middle C of the GCC triplet should be recognised by +3 Asp, or Glu, or Ser, or Thr (see Table 2 above). Thus a three-finger peptide gene was assembled from 8 overlapping synthetic oligonucleotides which were annealed and ligated according to standard procedures and the ~300bp product purified from a 2% agarose gel. The gene for finger 1 contained a partial codon randomisation at position +3 which allowed for inclusion of each of the above amino acids (D, E, S & T) and also certain other residues which were in fact not predicted to be desirable (e.g. Asr). The synthetic oligonucleotides were designed to have SfiI and NotI overhangs when annealed. The ~300 bp fragment was ligated into SfiI/NotI -cut FdSN vector and the ligation mixture was electroporated into DH5 α cells. Phage were produced from these as previously described and a selection step carried out using the G12V sequence (also as described) to eliminate phage without insert and those phage of the library which bound poorly.

Following selection, a number of separate clones were isolated and phage produced from these were screened by ELISA for binding to the G12V ras sequence and discrimination against the wild-type ras sequence. A number of clones were able to do this, and sequencing of phage DNA later revealed that these fell into two categories, one of which had the amino acid Asn at the +3 randomised position, and another which had two other undesirable mutations.

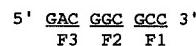
The appearance of Asn at position +3 is unexpected and most probably due to the fact that proteins with a cytosine-specific residue at position +3 bind to some *E. coli* DNA sequence so tightly that they are lethal. Thus phage display selection is not always guaranteed to produce the tightest-binding clone, since passage through bacteria is essential to the technique, and the selected proteins may be those which do not bind to the genome of this host if such binding is deleterious.

32

K_d measurements show that the clone with Asn+3 nevertheless binds the mutant G12V sequence with a K_d in the nM range and discriminates against the wild-type ras sequence. However it was predicted that Asn+3 should specify an adenine residue at the middle position, whereas the polypeptide we wished to make should specify a cytosine for optimal binding.

Thus we assembled a three-finger peptide with a Ser at position +3 of Finger 1 (as shown in FIG. 12), again for using synthetic oligos. This time the gene was ligated to pCANTAB5E phagemid. Transformants were isolated in the *E. coli* ABLE-C strain (from Stratagene) and grown at 30° C., which strain under these conditions reduces the copy number of plasmids so as to make their toxic products less abundant in the cells.

The amino acid sequence (Seq ID No. 18) of the fingers is shown in FIG. 12. The numbers refer to the α -helical 20 amino acid residues. The fingers (F1, F2 & F3) bind to the G12V mutant nucleotide sequence:



25 The bold A shows the single point mutation by which the G12V sequence differs from the wild type sequence.

Assay of the protein in eukaryotes (e.g. to drive CAT reporter production) requires the use of a weak oromoter. 30 When expression of the anti-RAS (G12V) protein is strong, the peptide presumably binds to the wild-type ras allele (which is required) leading to cell death. For this reason, a regulatable promoter (e.g. for tetracycline) will be used to deliver the protein in therapeutic applications, so that the 35 intracellular concentration of the protein exceeds the K_d for the G12V point mutated gene but not the K_d for the wild-type allele. Since the G12V mutation is a naturally occurring genomic mutation (not only a cDNA mutation as was the p190 bcr-abl) human cell lines and other animal 40 models can be used in research.

In addition to repressing the expression of the gene, the protein can be used to diagnose the precise point mutation present in the genomic DNA, or more likely in PCR amplified genomic DNA, without sequencing. It should therefore 45 be possible, without further inventive activity, to design diagnostic kits for detecting (e.g. point) mutations on DNA. ELISA-based methods should prove particularly suitable.

It is hoped to fuse the zinc finger binding polypeptide to an scFv fragment which binds to the human transferrin receptor, which should enhance delivery to and uptake by 50 human cells. The transferrin receptor is thought particularly useful but, in theory, any receptor molecule (preferably of high affinity) expressed on the surface of a human target cell could act as a suitable ligand, either for a specific immunoglobulin or fragment, or for the receptor's natural ligand fused or coupled with the zinc finger polypeptide.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 125

(2) INFORMATION FOR SEQ ID NO: 1:

-continued

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTCCTGCAGT TGGACCTGTG CCATGGCCGG CTGGGCCGCA TAGAATGGAA	50
CAACTAAAGC	60

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 92 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro	
5	10

Val Glu Ser Cys Asp Arg Arg Phe Ser Arg	
15	20

Ser Asp Glu Leu Thr Arg His Ile Arg Ile	
25	30

His Thr Gly Gln Lys Pro Phe Gln Cys Arg	
35	40

Ile Cys Met Arg Asn Phe Ser Xaa Xaa Xaa	
45	50

Xaa Leu Xaa Xaa His Xaa Arg Thr His Thr	
55	60

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys	
65	70

Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg	
75	80

Lys Arg His Thr Lys Ile His Leu Arg Gln	
85	90

Lys Asp

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TATGACTTGG ATGGGAGACC GCCTGG	26
------------------------------	----

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AATTCCAGGC GGTCTCCCAT CCAAGTCA	28
--------------------------------	----

6,013,453

35

36

-continued

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TATATAGCGT GGGCGTATAT A

21

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGTATATAAC GCCCACGCTA TATA

24

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TATATAGCGN NNGCGTATAT A

21

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCGTATATAAC GCNNNCGCTA TATA

24

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTCCATGGAG ACGCAGAACG CCTTCAGCGG CCA

33

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTCCATGGAG ACGCAGGGTGA GTTCCTCACG CCA

33

-continued

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCCCCTTCTC TTCCAGAAGC CCTTCAGCGG CCA

33

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
5 10

Ile Cys Met Arg Asn Phe Ser Asp Arg Ser
15 20

Ser Leu Thr Arg His Thr Arg His Thr Gly
25 30

Glu Lys Pro

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
5 10

Ile Cys Met Arg Asn Phe Ser Glu Arg Gly
15 20

Thr Leu Ala Arg His Glu Lys His Thr Gly
25 30

Glu Lys Pro

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
5 10

Ser Gln Gly Gly Asn Leu Val Arg His Leu
15 20

Arg His Thr Gly Glu Lys Pro

-continued

25

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
 5 10

Ser Gln Ala Gln Thr Leu Gln Arg His Leu
 15 20

Lys His Thr Gly Glu Lys
 25

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
 5 10

Ser Gln Ala Ala Thr Leu Gln Arg His Leu
 15 20

Lys His Thr Gly Glu Lys
 25

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
 5 10

Ser Gln Ala Gln Asp Leu Gln Arg His Leu
 15 20

Lys His Thr Gly Glu Lys
 25

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

-continued

Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
5 10

Ile Cys Met Arg Asn Phe Ser Asp Arg Ser
15 20

Ser Leu Thr Arg His Thr Arg Thr His Thr
25 30

Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys
35 40

Met Arg Asn Phe Ser Asp Arg Ser His Leu
45 50

Thr Arg His Thr Arg Thr His Thr Gly Glu
55 60

Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
65 70

Asn Phe Ser Asp Arg Ser Asn Leu Thr Arg
75 80

His Thr Arg Thr His Thr Gly Glu Lys
85

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Arg Ser Asp His Leu Thr Thr His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Arg Val Asp Ala Leu Glu Ala His Arg Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Gln Arg Ala Ser Leu Ala Ser His Met Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Aⁿs Arg Asp Thr Leu Thr Arg His Ser Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Gln Lys Gly His Leu Thr Glu His Arg Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Gln Ser Val His Leu Gln Ser His Ser Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Arg Leu Asp Gly Leu Arg Thr His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Thr Pro Gly Asn Leu Thr Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Asn Gly Gly Asn Leu Gly Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Arg Ala Asp Ala Leu Met Val His Lys Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Asn Gln Ser Asn Leu Glu Arg His His Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Asp Arg Ser Asn Leu Glu Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Arg Ser Asp Thr Leu Lys Lys His Gly Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Gln Gln Ser Asn Leu Val Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Asn Gly Ala Asn Leu Glu Arg His Arg Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Arg Glu Asp Ala Leu Thr Ser His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Arg Gly Asp His Leu Lys Asp His Ile Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Arg Gly Pro Asp Leu Ala Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Arg Glu Asp Val Leu Ile Arg His Gly Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Arg Ser Asp Leu Leu Gln Arg His His Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

Arg Gln Asp Thr Leu Val Gly His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

Arg Ala Ala Asp Leu Asn Arg His Val Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

Ser Gln Gly Asn Leu Gln Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 42:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

Thr Gly Gly Ser Leu Ala Arg His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Asp His Ala Asn Leu Ala Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

Leu Gln Ser Asn Leu Val Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 45:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

Arg Lys Asp Val Leu Val Ser His Val Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 46:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

Arg Arg Asp Val Leu Met Asn His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 47:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

Gln Gly Gly Asn Leu Val Arg His Leu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 48:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

Ser Arg Asp Val Leu Arg Arg His Asn Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

Glu Lys Ala Thr Leu Ala Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

Gln Ala Gln Thr Leu Gln Arg His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

Ile Ala Ser Asn Leu Leu Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

Arg Gly Asp His Leu Lys Asp His Ile Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

Arg Ser Asp His Leu Thr Thr His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 54:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

Gln Leu Ala His Leu Ser Thr His Lys Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 55:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Gln Ser Val His Leu Gln Ser His Ser Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

Gln Lys Gly His Leu Thr Glu His Arg Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

Gln Gly Gly Asn Leu Val Arg His Leu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Asn Gly Gly Asn Leu Gly Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Ala Arg Ser Asn Leu Leu Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 60:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

Leu Gln Ser Asn Leu Val Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Ile Ala Ser Asn Leu Leu Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 62:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

Asp Arg Ser Asn Leu Glu Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

Asn Gln Ser Asn Leu Glu Arg His His Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

Gln Gln Ser Asn Leu Val Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

Asn Gly Gly Asn Leu Gly Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 66:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

Asn Gly Ala Asn Leu Glu Arg His Arg Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 67:

- (i) SEQUENCE CHARACTERISTICS:

6,013,453

61

62

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Ser Gln Gly Asn Leu Gln Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 68:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

Ser His Pro Asn Leu Asn Arg His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 69:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

Thr Pro Gly Asn Leu Thr Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 70:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

Asp Arg Ser Asn Leu Glu Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 71:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

Asp His Ala Asn Leu Ala Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

Asp Arg Ser Ser Leu Thr Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 73:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

Glu Arg Gly Thr Leu Ala Arg His Glu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 74:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

Asp Arg Arg Leu Leu Asp Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 75:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

Asp Arg Ser Ser Leu Thr Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 76:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

Glu Arg Thr Ser Leu Ser Arg His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 77:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

Ser Ala Gly Thr Leu Val Arg His Ser Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 78:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

Gln Ala Gln Thr Leu Gln Arg His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 79:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

Glu Lys Ala Thr Leu Ala Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 80:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

Thr Gly Gly Ser Leu Ala Arg His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 81:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

Arg Gln Ser Thr Leu Gly Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 82:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

Glu Lys Ala Thr Leu Ala Arg His Met Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 83:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

Gln Ala Gln Thr Leu Gln Arg His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 84:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

Glu Arg Gly Thr Leu Ala Arg His Glu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 85:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

Gly Arg Asp Ala Leu Ala Arg His Gln Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 86:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

Arg Gly Pro Asp Leu Ala Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 87:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

Ser Arg Asp Val Leu Arg Arg His Asn Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 88:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

Arg Arg Asp Val Leu Met Asn His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 89:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

Arg Lys Asp Val Leu Val Ser His Val Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 90:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

Arg Arg Asp Val Leu Met Asn His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 91:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

Arg Glu Asp Ala Leu Thr Ser His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 92:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

Arg Val Asp Ala Leu Glu Ala His Arg Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 93:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

Asp Arg Ser Ser Leu Thr Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 94:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

Glu Arg Thr Ser Leu Ser Arg His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 95:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

Gly Ala Arg Ser Leu Thr Arg His Gln Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 96:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

Thr Gly Gly Ser Leu Ala Arg His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 97:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

Gln Arg Ala Ser Leu Ala Ser His Met Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 98:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

Asn Arg Asp Thr Leu Thr Arg His Ser Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 99:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

Glu Arg Gly Thr Leu Ala Arg His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 100:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

Arg Gly Asp Ala Leu Thr Ser His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 101:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

Arg Ala Asp Ala Leu Met Val His Lys Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 102:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

Arg Gln Asp Thr Leu Val Gly His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 103:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

Arg Gln Ser Thr Leu Val Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 104:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

Arg Ala Ala Asp Leu Asn Arg His Val Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 105:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

Arg Lys Asp Val Leu Val Ser His Val Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 106:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

Arg Arg Asp Val Leu Met Asn His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 107:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

Arg Ser Asp Thr Leu Lys Lys His Gly Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 108:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

Arg Gly Pro Asp Leu Ala Arg His Gly Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 109:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

Ala Arg Glu Val Leu Gln Arg His Thr Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 110:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

Arg Glu Asp Val Leu Ile Arg His Gly Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 111:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

Arg Ser Asp Leu Leu Gln Arg His His Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 112:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

Arg Leu Asp Gly Leu Arg Thr His Leu Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 113:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

Arg Gly Asp Ala Leu Thr Ser His Glu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 114:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

Arg Ala Asp Ala Leu Met Val His Lys Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 115:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

Arg Val Asp Ala Leu Glu Ala His Arg Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 116:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

Arg Arg Asp Val Leu Leu Asn His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 117:

- (i) SEQUENCE CHARACTERISTICS:

-continued

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

Arg Glu Asp Val Leu Ile Arg His Gly Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 118:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

Arg Ser Asp Leu Leu Gln Arg His His Lys
5 10

(2) INFORMATION FOR SEQ ID NO: 119:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

Lys Ala Ser Asn Leu Val Ser His Ile Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 120:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

Leu Arg His Asn Leu Glu Thr His Met Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 121:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

CGCAGAACCC

10

(2) INFORMATION FOR SEQ ID NO: 122:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

TCCAGAACGCC

10

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

CGCAGGTGAG

10

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

Glu	Gln	Lys	Leu	Ile	Ser	Glu	Glu	Asp	Leu	Asn
5									10	

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

Pro Lys Lys Lys Arg Lys Val

We claim:

1. A library of DNA sequences, each sequence encoding 45 a zinc finger polypeptide for display on a viral particle, the zinc finger polypeptide comprising at least three zinc fingers, with one zinc finger having a partially randomized allocation of amino acids being positioned between two or more zinc fingers having a defined amino acid sequence, the 50 partially randomized zinc finger having random allocation of amino acids at positions -1, +2, +3 and +6, position +1 being the first amino acid in the α -helix of the zinc finger.

2. A library according to claim 1, wherein the partially randomized zinc finger further has random allocation of 55 amino acids in at least one of the positions +1, +5 and +8.

3. A library according to claim 1 or 2, wherein the encoded partially randomized zinc finger is a zinc finger of the Zif268 polypeptide.

4. A library according to claim 1 as a fusion with the DNA 60 encoding the minor coat protein of bacteriophage fd.

5. A method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising the steps of:

comparing the binding to one or more DNA triplets of 65 each of a plurality of zinc finger polypeptides having a partially randomized zinc finger positioned between

two or more zinc fingers having defined amino acid sequence, the zinc finger polypeptides being encoded by a library in accordance with claim 1; and selecting those nucleic acid sequences encoding randomized zinc fingers which bind to the target DNA sequence.

6. A method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising the steps of:

screening against at least a portion of the target DNA sequence a plurality of zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence, the portion of the target DNA sequence being sufficient to allow binding of some of the zinc finger polypeptides, the zinc finger polypeptides being encoded by a library in accordance with claim 1;

comparing the binding to one or more DNA triplets of each of plurality zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence; and

selecting those nucleic acid sequences encoding randomized zinc fingers which bind to the target DNA sequence.

7. A method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, the method comprising the steps of:

screening against at least a portion of the target DNA sequence a plurality of zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence, the portion of the target DNA sequence being sufficient to allow binding of some of the zinc finger polypeptides, the zinc finger polypeptides being encoded by a library in accordance with claim 1;

comparing the binding to one or more DNA triplets of each of a plurality zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence;

selecting certain of the screened randomized zinc fingers for analysis of preferred binding characteristics;

and combining those sequences encoding desired zinc fingers to form a sequence encoding a zinc finger polypeptide which binds to the target DNA sequence.

8. A method for producing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising the steps of:

screening against at least a portion of the target DNA sequence a plurality of zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence, the portion of the target DNA sequence being sufficient to allow binding of some of the zinc finger polypeptides, the plurality of zinc finger polypeptides being coded by a library in accordance with claim 1;

selecting those nucleic acid sequences encoding randomized zinc fingers which bind to the target DNA sequence; and

expressing the selected nucleic acid sequences to produce zinc finger polypeptides which bind to the target DNA sequence.

9. A method of designing a zinc finger polypeptide for binding to a particular target DNA sequence, comprising the steps of:

screening against at least a portion of the target DNA sequence a plurality of zinc finger polypeptides having a partially randomized zinc finger positioned between two or more zinc fingers having defined amino acid sequence, the portion of the target DNA sequence being sufficient to allow binding of some of the zinc finger polypeptides, the zinc finger polypeptides being encoded by a library in accordance with claim 1; and

selecting those nucleic acid sequences encoding randomized zinc fingers which bind to the target DNA sequence.

10. A method according to claim 9, wherein two or more rounds of screening are performed.

11. A method of designing a zinc finger polypeptide for binding to a particular DNA target sequence, wherein a plurality of sequences encoding individual zinc fingers selected by the method of claim 9 are randomly combined in the appropriate order to encode a plurality of zinc finger polypeptides, the zinc finger polypeptides being screened against the target sequence, that combination of zinc finger sequences encoding a zinc finger polypeptide which binds to the target DNA sequence selected for use.

12. A method of modifying a nucleic acid sequence of interest present in a sample mixture by binding thereto a zinc finger polypeptide, wherein the zinc finger polypeptide is

designed in accordance with claim 9, comprising contacting the sample mixture with zinc finger polypeptide having affinity for at least a portion of the sequence of interest, so as to allow the zinc finger polypeptide to bind specifically to the sequence of interest.

13. A method according to claim 12, further comprising the step of separating the zinc finger polypeptide (and nucleic acid sequences specifically bound thereto) from the rest of the sample.

14. A method according to claim 12, wherein the zinc finger polypeptide is bound to a solid phase support.

15. A method according to claim 12, wherein the presence of the zinc finger polypeptide bound to the sequence of interest is detected by the addition of one or more detection reagents.

16. A method according to claim 12, wherein the DNA sequence of interest is present in an acrylamide or agarose gel matrix, or is present on the surface of a membrane.

17. A DNA library according to claim 1, consisting of 64 sequences, each sequence comprising a different one of the 64 possible permutations of a DNA triplet, the library being arranged in twelve sub-libraries, wherein for any one sub-library one base in the triplet is defined and the other two bases are randomized.

18. A library according to claim 17 wherein the sequences are biotinylated.

19. A library according to claim 17, wherein the sequences are associated with separation means.

20. A library according to claim 19, wherein the separation means is selected from the group consisting of: a microtitre plate; magnetic or non-magnetic beads or particles capable of sedimentation; and an affinity chromatography column.

21. A kit for making a zinc finger polypeptide for binding to a DNA of interest, comprising:

a library of DNA sequences, each sequence encoding a zinc finger polypeptide for display on a viral particle, the zinc finger polypeptide comprising at least three zinc fingers, with one zinc finger having a partially randomized allocation of amino acids being positioned between two or more zinc fingers having a defined amino acid sequence, the partially randomized zinc finger having random allocation of amino acids at positions -1, +2, +3 and +6. position +1 being the first amino acid in the α -helix of the zinc finger;

a vector molecule suitable for accepting one or more sequences from the library; and instructions for use.

22. A kit according to claim 21, wherein the vector directs the expression of the cloned sequences as a zinc finger polypeptide.

23. A kit according to claim 21, wherein the vector is capable of directing the expression of the cloned sequences as a zinc finger polypeptide displayed on the surface of a viral particle.

24. A kit according to claims 21 A library of DNA sequences, each sequence encoding a zinc finger polypeptide for display on a viral particle, the zinc finger polypeptide comprising at least three zinc fingers, with one zinc finger having a partially randomized allocation of amino acids being positioned between two or more zinc fingers having a defined amino acid sequence, the partially randomized zinc finger having random allocation of amino acids at positions -1, +2, +3 and +6, position +1 being the first amino acid in the α -helix of the zinc finger, further comprising a library consisting of 64 sequences, each sequence comprising a different one of the 64 possible permutations of a DNA triplet, the library being arranged in twelve sub-libraries,

87

wherein for any one sub-library one base in the triplet is defined and the other two bases are randomized.

25. A kit according to claim **24**, further comprising appropriate buffer solutions and/or reagents for detection of bound zinc fingers.

88

26. A kit according to claim **25**, further comprising a vector suitable for accepting one or more sequences selected from the library of DNA sequences encoding zinc fingers.

* * * * *

Positive and negative regulation of endogenous genes by designed transcription factors

Roger R. Beerli, Birgit Dreier, and Carlos F. Barbas III*

The Skaggs Institute for Chemical Biology and the Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037

Communicated by Richard A. Lerner, The Scripps Research Institute, La Jolla, CA, December 16, 1999 (received for review November 23, 1999)

Gene regulation by imposed localization was studied by using designed zinc finger proteins that bind 18-bp DNA sequences in the 5' untranslated regions of the protooncogenes erbB-2 and erbB-3. Transcription factors were generated by fusion of the DNA-binding proteins to repression or activation domains. When introduced into cells these transcription factors acted as dominant repressors or activators of, respectively, endogenous erbB-2 or erbB-3 gene expression. Significantly, imposed regulation of the two genes was highly specific, despite the fact that the transcription factor binding sites targeted in erbB-2 and erbB-3 share 15 of 18 nucleotides. Regulation of erbB-2 gene expression was observed in cells derived from several species that conserve the DNA target sequence. Repression of erbB-2 in SKBR3 breast cancer cells inhibited cell-cycle progression by inducing a G₁ accumulation, suggesting the potential of designed transcription factors for cancer gene therapy. These results demonstrate the willful up- and down-regulation of endogenous genes, and provide an additional means to alter biological systems.

The ability to specifically manipulate the expression of endogenous genes would have wide-ranging applications in medicine and in experimental and applied biology. To accomplish this, a number of promising approaches that aim to control gene expression have been described, operating either at the transcriptional level, such as polyamides, or the posttranscriptional level, such as antisense and ribozymes (1–3). While each of these methods may be applied advantageously in certain circumstances, they are not readily adapted to both gene activation and repression. Nature's control mechanisms center around transcription factors that function to direct the localization of enzymes to specific DNA addresses (4). Exploiting this fundamental principle for imposed control of gene expression is critically dependent on the availability of sequence-specific DNA-binding domains, the design of which has been the subject of intense research for many years. Of the DNA-binding motifs that have been studied, the modular zinc finger DNA-binding domains of the Cys₂-His₂ type have shown the most promise for the development of a universal system for gene regulation. Design studies and phage-based selections have shown that this motif is adaptable to the recognition of a wide variety of DNA sequences, often with exquisite specificity (5–10). Recently, we described a family of zinc finger domains that is sufficient for the construction of 17 million novel proteins that bind the 5'-(GNN)₆-3' family of DNA sequences. Further, we showed that these domains are functionally modular and may be recombined with one another to create polydactyl proteins capable of binding 18-bp sequences with the potential for genome-specific addressing (11, 12). While our early experiments have focused on the regulation of genes transiently introduced into cells, we realized that the willful and specific regulation of endogenous genes with designed transcription factors has remained an unmet challenge in biology.

Herein we target the endogenous erbB-2 and erbB-3 genes for imposed regulation. Both genes have been shown to be involved in human cancers (13–15). In particular, the erbB-2 gene is frequently overexpressed in human cancers, especially breast and ovarian, and elevated ErbB-2 levels correlate with a poor

prognosis (13). Moreover, there is increasing evidence that ErbB-3 is also involved in cancer, presumably by acting cooperatively with ErbB-2 (14–16). We show that transcription factors designed to bind in the transcribed regions of either erbB-2 or erbB-3 are capable of selectively up- or down-regulating expression of their respective target gene. The results presented herein demonstrate the targeted regulation of endogenous gene expression by using designed transcription factors and provide a foundation for wide-ranging applications of this technology.

Materials and Methods

Antibodies. The ErbB-2-specific antiserum 21N and mAb FSP77 were a gift from Nancy E. Hynes (17, 18). The ErbB-3-specific mAb SGP1 was from Oncogene Research Products. The ErbB-1-specific mAb EGFR1 and the phosphotyrosine-specific mAb PY20 were from Santa Cruz Biotechnologies. Fluorescently labeled, affinity-purified donkey F(ab')₂ anti-mouse IgG secondary antibodies were purchased from Jackson ImmunoResearch.

Expression Vectors. For doxycycline (Dox)-inducible expression, the E2C-KRAB and E2C-VP64 coding regions were PCR amplified from pcDNA3-based expression plasmids (12) and subcloned into pRevTRE (CLONTECH) by using BamHI and ClaI restriction sites (KRAB, Krüppel-associated box; VP64, tetrameric repeat of herpes simplex VP16's minimal activation domain). Fidelity of the PCR amplification was confirmed by sequencing. For retroviral expression of the E2C and E3 proteins, the E2C-KRAB and E2C-VP64 coding regions were cloned into pMX-IRES-GFP (19) by using BamHI and NotI restriction sites (IRES, internal ribosome-entry site; GFP, green fluorescent protein). The E2C coding region was then exchanged for the E3 coding region by SfiI digestion. For transient expression in reporter assays, the various E2C 3 finger-VP64 fusion constructs were assembled in pcDNA3 as described (12).

Retroviral Gene Targeting. The retroviral pMX-IRES-GFP/zinc finger constructs were transiently transfected into the amphotropic packaging cell line Phoenix Amphi by using Lipofectamine Plus (GIBCO/BRL) and, 2 days later, culture supernatants were used for infection of target cells in the presence of 8 µg/ml Polybrene. Three days after infection, cells were harvested for analysis.

Flow Cytometric Analysis. Cells were trypsinized and washed in fluorescence-activated cell sorting (FACS) buffer [phosphate-

Abbreviations: Dox, doxycycline; UTR, untranslated region; KRAB, Krüppel-associated box; VP64, tetrameric repeat of herpes simplex VP16's minimal activation domain; IRES, internal ribosome-entry site; GFP, green fluorescent protein.

*To whom reprint requests should be addressed at: The Scripps Research Institute, BCC-515, North Torrey Pines Road, La Jolla, CA 92037. E-mail: carlos@scripps.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: Proc. Natl. Acad. Sci. USA, 10.1073/pnas.040552697. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.040552697

buffered saline (PBS) containing 0.1% sodium azide and 1% BSA] prior to staining. Approximately 10^6 cells were then resuspended in 100 μ l of FACS buffer containing 5 μ g/ml of the respective mouse mAb. After incubation on ice for 1 h, cells were washed twice in FACS buffer. Bound antibodies were stained with fluorescently labeled anti-mouse secondary antibody. Finally, the cells were washed twice in FACS buffer, resuspended in 500 μ l of FACS buffer, and analyzed for their fluorescence with a Becton Dickinson FACScan.

Luciferase Assays. These assays were performed as described (12).

Construction and Characterization of E3 Protein. For the construction of the E3 six-finger protein, DNA recognition helices from the Zif268 finger 2 variants pmGGA, pGCC, and pGTC were utilized (10). Initially, two three-finger proteins binding each of the 9-bp half-sites of the 18-bp target sequence were constructed, by grafting the appropriate DNA recognition helices into the framework of the three-finger protein Sp1C. DNA fragments encoding the two three-finger proteins were assembled from six overlapping oligonucleotides as described (12). The six-finger protein coding region was then assembled in the bacterial expression vector pMal-CSS. E3 and E2C zinc finger proteins were expressed in the *Escherichia coli* strain XL1-Blue, purified, and analyzed by ELISA and electrophoretic mobility-shift assays as described (10).

Generation of Stable HeLa cell Clones. The pRevTRE/E2C-KRAB and pRevTRE/E2C-VP64 constructs were transfected into the HeLa/tet-off cell line (20) by using Lipofectamine Plus reagent. After 2 weeks of selection in hygromycin-containing medium, in the presence of 2 μ g/ml Dox, stable clones were isolated and analyzed for Dox-dependent regulation of ErbB-2 expression.

Western Blotting, Immunoprecipitations, and Northern Blotting. These procedures were carried out essentially as described (21).

Cell Cycle Analysis. Retrovirus-infected cells were stained with mAb FSP77 in combination with fluorescein-conjugated secondary antibody as described above. Cells were then fixed and permeabilized in PBS/4% paraformaldehyde/0.1% saponin for 10 min at room temperature. After two washes with PBS/0.1% saponin, cells were stained with 7-aminoactinomycin (5 μ g/ml in PBS/0.1% saponin) for 30 min on ice. After two more washes in PBS/0.1% saponin, cells were subjected to flow cytometric analysis. The fluorescence of the retrovirus-encoded GFP, measured in the same channel as fluorescein, was relatively weak and did not interfere with the ErbB-2 staining.

Results and Discussion

Imposed Transcriptional Regulation of the Endogenous erbB-2 Gene. We have previously described the generation of designed transcription factors capable of specifically regulating an erbB-2 promoter-luciferase reporter construct (12). Here we target the endogenous erbB-2 gene for imposed regulation. Endogenous genes are packaged within chromatin and are controlled by a multiplicity of cis- and trans-acting factors (22, 23), making it not known *a priori* whether specific gene regulation imposed with a dominant designed transcription factor is possible.

Thus, the potent transcriptional repressor E2C-KRAB and the transactivator protein E2C-VP64 (12) were tested for their ability to impose a dominant regulatory effect on the endogenous erbB-2 gene. Both proteins contain the same designed zinc finger protein, E2C, that recognizes the 18-bp DNA sequence 5'-GGG GCC GGA GCC GCA GTG-3' in the 5' untranslated region (UTR) of the protooncogene erbB-2. This DNA-binding protein was constructed from six Sp1C-based, predefined, and modular zinc finger domains, each recognizing a specific 3-bp

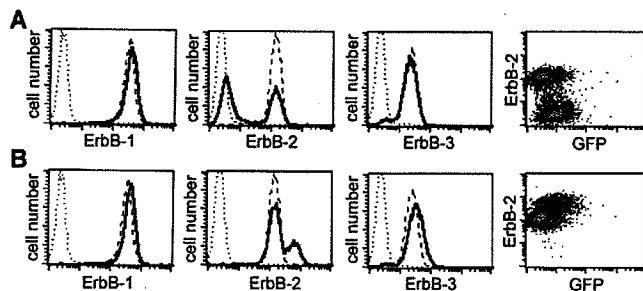


Fig. 1. Retrovirus-mediated erbB-2 gene targeting. A431 cells were infected with E2C-KRAB- (A) or E2C-VP64- (B) encoding retrovirus. Three days later, intact cells were stained with the ErbB-1-specific mAb EGFR1, the ErbB-2-specific mAb FSP77, or the ErbB-3-specific mAb SGP1 in combination with phycoerythrin-labeled secondary antibody, and analyzed by flow cytometry. Dotted lines, control staining (primary antibody omitted) of mock-infected cells; dashed lines, specific staining of mock-infected cells; solid lines, specific staining of, respectively, E2C-KRAB- or E2C-VP64-infected cells.

subsite (10, 12). The repressor protein contains the Kox-1 KRAB domain (24), whereas the transactivator VP64 contains a tetrameric repeat of the minimal activation domain (25) derived from the herpes simplex virus protein VP16.

For high-efficiency expression of the E2C-KRAB and E2C-VP64 proteins in various cell lines, their coding regions were introduced into the retroviral vector pMX-IRES-GFP (19). This vector expresses a single bicistronic message for the translation of the zinc finger protein and, from an IRES, the GFP. Because both coding regions share the same mRNA, their expression is physically linked to one another and GFP expression is an indicator of zinc finger expression. Virus prepared from these plasmids was then used to infect the human carcinoma cell line A431. Three days after infection, ErbB-2 expression was measured by flow cytometry. Significantly, about 59% of the E2C-KRAB virus-treated cells were essentially ErbB-2 negative, whereas in about 27% of the E2C-VP64 virus-treated cells ErbB-2 levels were increased (Fig. 1). Plotting of GFP fluorescence vs. ErbB-2 fluorescence revealed that there were two cell populations, one with normal ErbB-2 levels that was GFP negative, and another with altered ErbB-2 levels that was GFP positive (Fig. 1). Specificity of gene targeting was investigated by measuring the expression levels of the related ErbB-1 and ErbB-3 proteins. No significant alterations of the levels of these proteins were detected, indicating that erbB-2 gene targeting is specific and not a nonspecific result of general alterations in gene expression or overexpression of the effector domains (Fig. 1). The lack of any appreciable regulation of erbB-3 is particularly remarkable because its 5'-UTR contains the 18-bp sequence 5'-GGA GCC GGA GCC GCA GTG-3' (26, 27), which presents only three mismatches to E2C's designed target sequence (15-bp identity—lowercase letters indicate differences).

Six-Finger Proteins Are Required for Efficient Imposed Regulation. To evaluate whether targeting 18 bp of DNA sequence with a six-finger protein is necessary for efficient target gene regulation, three-finger proteins were also analyzed. Thus, the two three-finger constituents of the E2C protein, binding either of the 9-bp half-sites of the E2C target sequence, were converted into transcriptional activators by fusion with the synthetic transactivation domain VP64. Significantly, while the E2C-VP64 six-finger fusion protein efficiently activated transcription of an erbB-2-luciferase reporter, none of the three-finger fusion proteins had an appreciable effect (Fig. 2). The E2C six-finger protein has a roughly 50-fold higher affinity for the E2C target site than its constituent three-finger proteins, which bind their target DNA with dissociation constants of 25 and 35 nM (12).

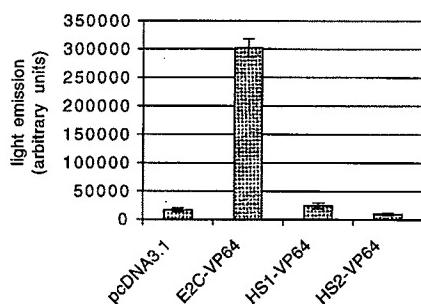


Fig. 2. Luciferase reporter gene assay. HeLa cells were cotransfected with the indicated zinc finger expression plasmids and an *erbB-2* promoter (-758 to -1)-luciferase reporter construct. Luciferase activity in total cell extracts was measured 48 h after transfection. Each bar represents the mean value (\pm standard deviation) of duplicate measurements. HS1 and HS2, three-finger proteins binding, respectively, half-site 1 or half-site 2 of the 18-bp E2C target sequence. pcDNA3.1 is a control plasmid that does not express a transcription factor.

Thus, these results suggest that not only the increased specificity but also the significantly higher affinity of six-finger proteins may be required to provide a dominant effect on gene regulation.

Regulation of erbB-2 Expression in Nonhuman Cells. The zinc finger target sequence within erbB-2's 5'-UTR lies within a 28-bp sequence stretch that is conserved in many species (28). To investigate regulation of erbB-2 gene expression in nonhuman primate cells, African green monkey COS-7 fibroblasts were infected with the bicistronic E2C-KRAB retrovirus and analyzed by flow cytometry. As in human cells, expression of the repressor protein as indicated by the GFP marker correlated well with a loss of ErbB-2 protein (Fig. 3A). Similarly, gene targeting in murine cells was evaluated by infection of NIH 3T3 cells with E2C-KRAB- and E2C-VP64-encoding retrovirus. ErbB-2 expression levels were then monitored by Western blotting rather than flow cytometry, because of a lack of reactivity of the mAb with the murine ErbB-2 extracellular domain. Here again, with E2C-KRAB a complete transcriptional knockout upon correction for infected cells was observed (Fig. 3B). However, unlike in human cell lines, E2C-VP64-induced ErbB-2 up-regulation was rather modest in NIH 3T3 cells, approximately 1.8-fold upon correction for infection efficiency (Fig. 3B). A likely explanation for this discrepancy lies in the different structures of the human and mouse promoters. The mouse erbB-2 promoter, unlike the human, does not contain a TATA box (28). Transcriptional

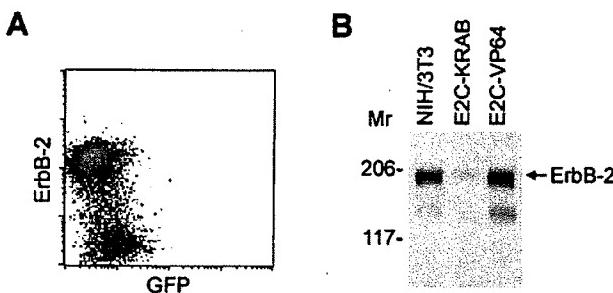


Fig. 3. erbB-2 gene targeting in nonhuman cells. (A) Flow cytometric analysis of ErbB-2 expression. COS-7 fibroblasts were infected with E2C-KRAB-encoding retrovirus. Three days later, intact cells were stained with the ErbB-2 specific mAb FSP77 in combination with phycoerythrin-labeled secondary antibody, and analyzed by flow cytometry. (B) ErbB-2 Western blot. NIH 3T3 fibroblasts were mock-infected or infected with E2C-KRAB- or E2C-VP64-encoding retrovirus. Three days later, protein extracts were prepared and subjected to Western blotting with the ErbB-2-specific antiserum 21N. Mr, molecular weight $\times 10^{-3}$.

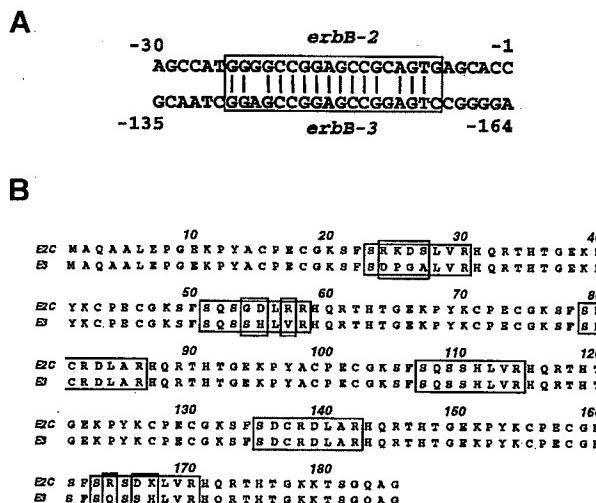


Fig. 4. (A) Alignment of E2C target sequence in the erbB-2 5'-UTR with the E3 target sequence in the erbB-3 5'-UTR. Numbers indicate the distance from the ATG translation initiation codon. (B) Amino acid sequence alignment of E2C and E3 proteins. DNA recognition helix sequence positions –1 to 6 of each finger, as well as sequence differences, are boxed.

activation by VP16 is, at least in part, mediated by its interaction with TFIID, a multiprotein complex also containing the TATA-binding protein (29). It is therefore plausible that the E2C-VP64 protein activates transcription less effectively in the absence of a TATA box. These data show that while a DNA-binding site may be conserved with respect to sequence and relative position within a target cell, effector domains may need to be optimized for maximal efficiency due to context effects. Nevertheless, while their potencies may differ, the designed transcription factors described here are capable of imposing regulation of *erbB-2* gene transcription in cells derived from different species, providing a strategy for the study of gene function in a variety of organisms.

Construction and Characterization of a Polydactyl Protein for Regulation of the *erbB-3* Gene. Examination of the *erbB-3* 5'-UTR revealed the presence of an 18-bp sequence that was highly similar to the E2C target sequence in the *erbB-2* 5'-UTR (26,27). Although they are at different distances and orientations with respect to the ATG initiation codons, the two sequences differ by only three nucleotides (Fig. 4A). Thus, we decided to construct a six-finger protein recognizing this sequence to investigate whether transcription factors could be designed to selectively regulate *erbB-3* gene expression.

We have previously described several strategies for the construction of polydactyl proteins from defined, modular building blocks (10, 12). The most successful strategy involved grafting of the amino acid residues of each zinc finger involved in base-specific DNA recognition (a short α -helical region referred to as the "recognition helix") into the framework of the designed consensus protein Sp1C, a derivative of the transcription factor Sp1 (30). Thus, the six-finger protein E3 designed to bind the 18-bp erbB-3 target sequence was built by using the Sp1C helix grafting strategy, the same method used for construction of the E2C protein described herein. An alignment of the E2C and E3 proteins reveals extensive sequence identity (Fig. 4B). In particular, the entire protein framework, as well as three of the six recognition helices, are identical. Only the recognition helices of fingers 1, 2, and 6 were partially different, reflecting the fact that the 3-bp subsites recognized by these fingers differed by 1 nucleotide each.

For a detailed analysis of its binding properties, the E3 protein

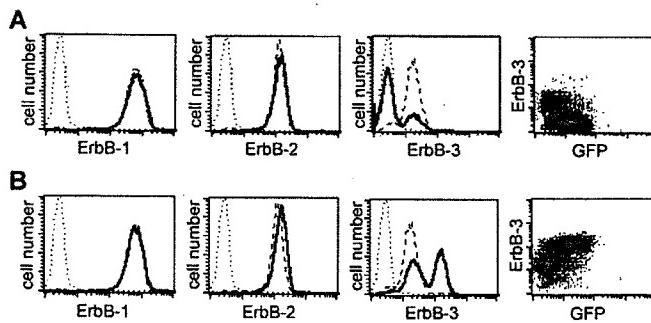


Fig. 5. Retrovirus-mediated *erbB-3* gene targeting. A431 cells were infected with E3-KRAB- (**A**) or E3-VP64- (**B**) encoding retrovirus. Three days later, intact cells were stained with the ErbB-1 specific mAb EGFR1, the ErbB-2 specific mAb FSP77, or the ErbB-3 specific mAb SG1 in combination with phycoerythrin-labeled secondary antibody, and analyzed by flow cytometry. Dotted lines, control staining (primary antibody omitted) of mock-infected cells; dashed lines, specific staining of mock-infected cells; solid lines, specific staining of, respectively, E3-KRAB- or E3-VP64-infected cells.

was purified as a fusion with the maltose-binding protein. Initially, an ELISA analysis was carried out, revealing specific binding of the E3 protein to its target site, with little or no crossreactivity to various other 5'-*(GNN)₆-3'* DNA sequences (not shown). A similar observation was made with the E2C protein (12). However, because of the similarity of the DNA sequences recognized, some crossreactivity of the two proteins with each other's target site was detected (not shown). To obtain a quantitative measure for the extent of discrimination between target and nontarget sequence, the affinities of the two proteins to each target sequence was determined by electrophoretic mobility-shift assay (10). These studies revealed high-affinity binding of the E3 protein to its target, with a K_d value of 0.35 nM ($\pm 10\%$), whereas the affinity of binding to the E2C target sequence was about 30-fold lower, with a K_d value of 10 nM ($\pm 15\%$). Similarly, the affinity of the E2C protein to its target was subnanomolar, with a K_d value of 0.75 nM ($\pm 15\%$) as we reported previously (12), whereas binding to the E3 site was significantly weaker, with a K_d value of 11 nM ($\pm 30\%$). Thus, both the E2C and the E3 proteins bind their respective target sequence with very high affinity and are able to discriminate between their cognate and very closely related DNA sequences.

Imposed Transcriptional Regulation of the Endogenous *erbB-3* Gene. Designed transcription factors were generated by fusing the E3 protein to repression or activation domains. In a manner analogous to the E2C fusion constructs, the E3-KRAB protein was produced by fusing the KRAB repressor domain to E3's N terminus, while E3-VP64 was generated by fusing the synthetic VP64 transactivation domain to its C terminus.

To analyze the ability of the *erbB-3*-specific transcription factors to impose a dominant regulatory effect on the native *erbB-3* gene, the E3-KRAB and E3-VP64 coding regions were introduced into the retroviral vector pMX-IRES-GFP. Retroviruses prepared from this vector were then used to infect A431 cells. Three days after infection, expression levels of various members of the ErbB receptor family were monitored by flow cytometry. Dramatic alterations in the levels of ErbB-3 were detected in significant fractions of infected cell populations. Expression was abolished in 74% of E3-KRAB virus-infected cells, whereas almost 8-fold higher ErbB-3 levels were detected in 48% of E3-VP64 virus-infected cells. Plotting of ErbB-3 fluorescence against GFP fluorescence revealed that only GFP-positive, i.e., infected, cells displayed altered ErbB-3 levels (Fig. 5). Thus, E3-based transcription factors are as potent as E2C-based transcription factors in regulating target gene expression.

In contrast to the efficient regulation of ErbB-3 expression, neither E3-KRAB nor E3-VP64 significantly affected ErbB-1 and ErbB-2 expression levels (Fig. 5). Given the similarity of the E3 and E2C target sequences, the lack of a significant effect on *erbB-2* gene expression is yet another demonstration of the exquisite specificity inherent to the zinc finger-based gene switches described here.

Requirements for Imposing Specific Regulation on Endogenous Genes.

The extent of discrimination between target and nontarget gene exhibited by the E2C- and E3-based fusion proteins has important implications for the future design of artificial gene switches. In particular, it is possible to make general predictions on the affinity with which a transcription factor has to bind to the promoter of a gene of interest to impose a dominant transcriptional control. We find that binding with K_d values of 10 nM or higher is not sufficient, as evidenced by the lack of *erbB-3* gene regulation by E2C fusion proteins (Fig. 1), the lack of *erbB-2* gene regulation by E3 fusion proteins (Fig. 5), and the lack of *erbB-2* promoter regulation by three-finger fusion proteins (Fig. 2). Binding with a K_d value of around 1 nM, however, appears to be associated with an occupancy of the target site sufficient for imposed gene control, as evidenced by the efficient control of target gene expression exhibited by the E2C- and E3-fusion proteins. Thus, proteins with significantly better affinities may be undesirable because increased binding to nonspecific DNA sequences may lead to significant side effects. Further, the length of the DNA sequence bound is also key to specific regulation. A three-finger protein binding 9 bp of DNA sequence would be expected to find more than 10^4 binding sites in the human genome, whereas an 18-bp binding site targeted by a six-finger protein has the potential to be unique within all genomes. Thus one could expect that any regulatory effect imposed by targeting only 9 bp of sequence would be nonspecific.

Placing Expression of an Endogenous Gene Under Chemical Control. In many circumstances, constitutive up- or down-regulation of a given gene of interest may not be desirable. To evaluate the feasibility of affecting target gene expression in an inducible manner, the E2C-KRAB and E2C-VP64 proteins were expressed from a regulatable promoter. For this purpose, a derivative of the human cervical carcinoma cell line HeLa, HeLa/tet-off, was utilized (20). Because HeLa cells are of epithelial origin they express ErbB-2 and are well suited for studies of *erbB-2* gene targeting. HeLa/tet-off cells produce the tetracycline-controlled transactivator, allowing induction of a gene of interest under the control of a tetracycline response element (TRE) by removal of tetracycline or its derivative doxycycline (Dox) from the growth medium. Thus, the pRevTRE/E2C-SKD and pRevTRE/E2C-VP64 plasmids were constructed and transfected into HeLa/tet-off cells, and 20 stable clones each were isolated and analyzed for Dox-dependent target gene regulation. As a read-out of *erbB-2* promoter activity, ErbB-2 protein levels were initially analyzed by Western blotting. A significant fraction of these clones showed regulation of ErbB-2 expression upon removal of Dox for 4 days—i.e., down-regulation of ErbB-2 in E2C-KRAB clones and up-regulation in E2C-VP64 clones (Fig. 6A). ErbB-2 protein levels were correlated with altered levels of their specific mRNA, indicating that regulation of ErbB-2 expression was a result of repression or activation of transcription (Fig. 6B). The additional ErbB-2 protein expressed in E2C-VP64 clones was indistinguishable from naturally expressed protein and biologically active, since epidermal growth factor (EGF) readily induced its tyrosine phosphorylation (Fig. 6C). The ErbB-2 levels in the E2C-KRAB clone 27, in the absence of Dox, were below the level of detection, as was its EGF-induced tyrosine phosphorylation. Therefore, ErbB-2 expression was also analyzed by flow cytometry, revealing no detectable ErbB-2

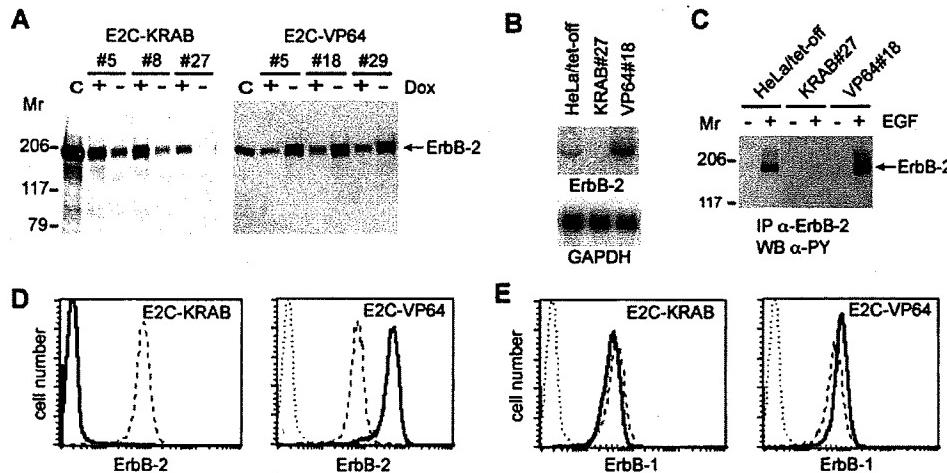


Fig. 6. *erbB-2* gene targeting in stable HeLa cell clones. (A) ErbB-2 Western blot. The indicated E2C-KRAB- and E2C-VP64-expressing clones were maintained in the presence or absence of 2 μ g/ml Dox for 4 days. Protein extracts from these cells were subjected to Western blotting with the ErbB-2 specific antiserum 21N. Lane C, HeLa/tet-off extract. (B) Northern blot. Total RNA extracted from the indicated cell lines maintained in the absence of Dox for 4 days was subjected to Northern blotting with an *erbB-2* specific probe. The membrane was stripped and reprobed with a glyceraldehyde-3-phosphate (GAPDH)-specific probe as a control. (C) Epidermal growth factor (EGF)-induced tyrosine phosphorylation of ErbB-2. The indicated cell lines were maintained in the absence of Dox for 4 days, serum starved overnight, and either induced with 100 ng/ml EGF for 10 min at room temperature or left untreated. ErbB-2 was immunoprecipitated (IP) from protein extracts with antiserum 21N and analyzed by Western blotting (WB) with mAb PY20. (D and E) Flow cytometric analysis of ErbB-2 and ErbB-1 expression. Cells were maintained for 4 days in the absence of Dox, stained with mAbs FSP77 or EGFR1 in combination with phycoerythrin-labeled secondary antibody, and analyzed for their fluorescence in a FACScan (Becton Dickinson). Dotted lines, control staining (primary antibody omitted) of HeLa/tet-off cells; dashed lines, specific stainings of HeLa/tet-off cells; solid lines, specific stainings of, respectively, Dox-deprived KRAB clone 27 and VP64 clone 18.

expression in E2C-KRAB clone 27, in sharp contrast to the dramatic up-regulation (5.6-fold) of ErbB-2 in E2C-VP64 clone 18 (Fig. 6D). Thus, the extent of *erbB-2* gene regulation ranged from total repression (E2C-KRAB clone 27) to almost 6-fold activation (E2C-VP64 clone 18). No significant effect on the expression of the related ErbB-1 protein was observed, indicating that regulation of ErbB-2 expression was not a result of general down- or up-regulation of transcription (Fig. 6E). In summary, these results show that it is feasible to use designed transcription factors to place the expression of an endogenous gene under control of an exogenous chemical inducer.

Toward a Therapeutic Application of Artificial Gene Switches. Over-expression of ErbB-2 leads to constitutive activation of its intrinsic tyrosine kinase activity (31), and it has been shown that down-regulation of ErbB-2 in tumor cells overexpressing the receptor leads to growth inhibition (32–34). The mechanism of growth inhibition appears to be that progression of the cells from the G₁ to the S phase of the cell cycle is prevented (35). Thus, we investigated whether expression of our designed transcriptional repressor in ErbB-2-overexpressing tumor cells would lead to a G₁ block. Therefore, SKBR3 breast cancer cells were infected with E2C-KRAB retrovirus and cell-cycle distribution was analyzed in relation to ErbB-2 expression levels by flow cytometry (Fig. 7A). Two cell populations were observed: about 40% of the cells were not infected and had normal ErbB-2 levels (Fig. 7A, Left, cell population 1), whereas the infected cells, ≈60%, displayed approximately 7-fold-reduced receptor levels after 3 days (Fig. 7A, Left, cell population 2). Compared with cells with normal receptor levels, a significantly larger fraction of cells with decreased ErbB-2 expression levels was in the G₁ phase of the cell cycle (Fig. 7A, Center and Right). To ascertain that the G₁ accumulation observed with SKBR3 cells was specific for ErbB-2-overexpressing tumor cells, a similar analysis was carried out with the T47D breast cancer cell line, which does not display elevated levels of ErbB-2 (Fig. 7B). Indeed, when T47D cells were infected with the E2C-KRAB retrovirus and subjected to flow cytometric analysis, cell populations with normal and

reduced ErbB-2 levels were found to display indistinguishable DNA contents. Thus, our designed repressor protein is able to specifically induce G₁ accumulation of ErbB-2-overexpressing tumor cells. The ability to inhibit cell-cycle progression, and hence inhibit growth of ErbB-2-overexpressing tumor cells, suggests the potential of designed transcription factors for cancer gene therapy.

Conclusions. In summary, artificial transcription factors can be designed to impose a dominant regulatory effect on the transcription of endogenous genes in their native chromosomal context, and if desired to place them under chemical control. We anticipate that this strategy will find applications in gene therapy and in basic and applied research where modulation of gene transcription can be employed to dissect biological mechanisms or to alter phenotypes of cells and organisms.

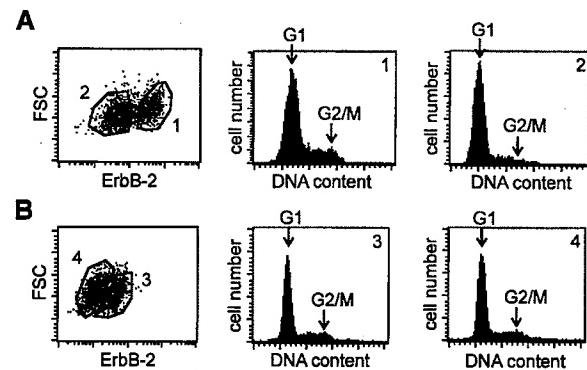


Fig. 7. Flow cytometric cell cycle analysis. SKBR3 (A) and T47D cells (B) were infected with E2C-KRAB-encoding retrovirus. Three days later, cells were stained with the ErbB-2-specific mAb FSP77 in combination with fluorescein-labeled secondary antibody (Left), as well as with 7-aminoactinomycin to show cell cycle distribution by DNA content (Center and Right), and analyzed by flow cytometry. Cell cycle histograms were generated from cells gated as indicated by numbers. FSC, forward scattered light.

We are very grateful to Nancy E. Hynes for providing us with antibodies, to Harvey Lodish for the pMX-IRES-GFP vector, and to Jody Berry for his critical reading of the manuscript. This study was supported by grants from the National Institutes of Health and the

Skaggs Institute for Chemical Biology to C.F.B. R.R.B. was the recipient of fellowships from the Krebsliga beider Basel and the Swiss Cancer League. B.D. was a recipient of a fellowship from the Deutsche Forschungsgemeinschaft.

1. Zaug, A. J., Been, M. D. & Cech, T. R. (1986) *Nature (London)* **324**, 429–433.
2. Matteucci, M. D. & Wagner, R. W. (1996) *Nature (London)* **384**, 20–22.
3. Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E. & Dervan, P. B. (1997) *Nature (London)* **387**, 202–205.
4. Ptashne, M. & Gann, A. (1998) *Curr. Biol.* **8**, 812–822.
5. Desjarlais, J. R. & Berg, J. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7345–7349.
6. Choo, Y. & Klug, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11163–11167.
7. Rebar, E. J. & Pabo, C. O. (1994) *Science* **263**, 671–673.
8. Wu, H., Yang, W.-P. & Barbás, C. F., III (1995) *Proc. Natl. Acad. Sci. USA* **92**, 344–348.
9. Greisman, H. A. & Pabo, C. O. (1997) *Science* **275**, 657–661.
10. Segal, D. J., Dreier, B., Beerli, R. R. & Barbas, C. F., III (1999) *Proc. Natl. Acad. Sci. USA* **96**, 2758–2763.
11. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F., III (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5525–5530.
12. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F., III (1998) *Proc. Natl. Acad. Sci. USA* **95**, 14628–14633.
13. Hynes, N. E. & Stern, D. F. (1994) *Biochim. Biophys. Acta* **1198**, 165–184.
14. Kraus, M. H., Fedi, P., Starks, V., Muraro, R. & Aaronson, S. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2900–2904.
15. Alimandi, M., Romano, A., Curia, M. C., Muraro, R., Fedi, P., Aaronson, S. A., Di Fiore, P. P. & Kraus, M. H. (1995) *Oncogene* **10**, 1813–1821.
16. Siegel, P. M., Ryan, E. D., Cardiff, R. D. & Muller, W. J. (1999) *EMBO J.* **18**, 2149–2164.
17. Hynes, N. E., Gerber, H. A., Saurer, S. & Groner, B. (1989) *J. Cell. Biochem.* **39**, 167–173.
18. Harwerth, I.-M., Wels, W., Marte, B. M. & Hynes, N. E. (1992) *J. Biol. Chem.* **267**, 15160–15167.
19. Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A. & Lodish, H. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 10669–10674.
20. Gossen, M. & Bujard, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5547–5551.
21. Graus-Porta, D., Beerli, R. R. & Hynes, N. E. (1995) *Mol. Cell. Biol.* **15**, 1182–1191.
22. Holstege, F. C. P. & Young, R. A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 2–4.
23. Chen, L. (1999) *Curr. Opin. Struct. Biol.* **9**, 48–55.
24. Margolin, J. F., Friedman, J. R., Meyer, W. K.-H., Vissing, H., Thiesen, H.-J. & Rauscher, F. J., III (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4509–4513.
25. Seipel, K., Georgiev, O. & Schaffner, W. (1992) *EMBO J.* **11**, 4961–4968.
26. Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. & Aaronson, S. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9193–9197.
27. Plowman, G. D., Whitney, G. S., Neubauer, M. G., Green, J. M., McDonald, V. L., Todaro, G. J. & Shoyab, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4905–4909.
28. White, M. R. & Hung, M. C. (1992) *Oncogene* **7**, 677–683.
29. Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J. & Greenblatt, J. (1991) *Nature (London)* **351**, 588–590.
30. Desjarlais, J. R. & Berg, J. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2256–2260.
31. Di Fiore, P. P., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R. & Aaronson, S. A. (1987) *Science* **237**, 178–182.
32. Hudziak, R. M., Lewis, G. D., Winget, M., Fendly, B. M., Shepard, H. M. & Ullrich, A. (1989) *Mol. Cell. Biol.* **9**, 1165–1172.
33. Deshane, J., Loechel, F., Conry, R. M., Siegel, G. P., King, C. R. & Curiel, D. T. (1994) *Gene Ther.* **1**, 332–337.
34. Daly, J. M., Jannot, C. B., Beerli, R. R., Graus-Porta, D., Maurer, F. G. & Hynes, N. E. (1997) *Cancer Res.* **57**, 3804–3811.
35. Neve, R. M., Sutterluty, H., Pullen, N., Lane, H. A., Daly, J. M., Krek, W. & Hynes, N. E. (2000) *Oncogene*, in press.

DNA-Binding Proteins Turn Genes On And Off

Scientists at Scripps Research Institute, La Jolla, Calif., have designed proteins that bind DNA in a highly specific manner and act like transcription factors to turn endogenous genes on and off in living cells. Transcription factors are DNA-binding proteins that activate the transcription of DNA into messenger RNA.

Postdoctoral fellows Roger R. Berni and Birgit Dreyer and molecular biology professor Carlos F. Barbas III report the findings in the *Proceedings of the National Academy of Sciences USA* [97, 1495 (2000)].

This is the first time we've been able to show that these designed transcription factors work on real genes and real chromosomes, not genes or binding sites that have been introduced into cells," Barbas says. "We can really knock out expression in a controlled way or elevate the level of expression from the native level to far beyond it to study the role of proteins within the cell." He notes that the ability of the designed transcription factors to bind 18-base-pair sequences makes it possible for each of them to target a unique site within the human genome.

Barbas says that his group's technology can potentially be used to regulate genes of pharmaceutical interest, attack cancer or genetic diseases, control crops, and even "make flowers bloom when you want them to." But he believes the technology also has great potential as a basic research tool. "One could make switches that turn on and turn off genes in a very high-throughput fashion and study the phenotypes of the cells that result," he says.

Chemistry professor Peter B. Dervan of California Institute of Technology calls the study by Barbas and coworkers "a step forward. Long term, designed transcription factor mimics are a whole new approach to human therapeutics," he says.

"There are several milestones one needs to overcome in the design and discovery of artificial transcription factors to control gene expression, and hence human disease," Dervan adds. For example, "putting artificial proteins in human cells will require gene therapy techniques. But that is a medical tech-

nology that will likely be successful in our lifetime."

The synthetic transcription factors developed by Barbas and coworkers are designed proteins containing DNA-binding elements called zinc fingers, structures found in some natural transcription factors. Zinc fingers—individually folding domains of about 30 amino acid residues centered on a zinc ion—were discovered in 1985 by molecular biologist Aaron Klug and coworkers at the Medical Research Council Laboratory of Molecular Biology, Cambridge, England. "We call them zinc fingers because they grip or grasp DNA," Klug says.

Natural transcription factors often have both a DNA-binding domain that localizes the protein to a specific site and one or more effector domains that activate or repress transcription at or near that site. In the designed transcription factors devised by Barbas' group, synthetic zinc finger DNA-binding do-

mains, which are designed to target specific sequences, are combined with naturally derived effector domains.

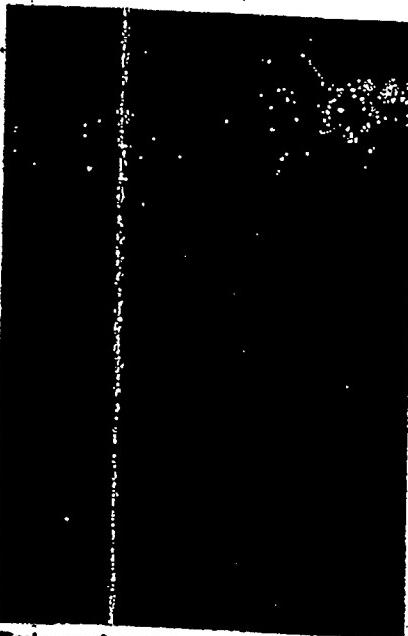
In their *PNAS* paper, Barbas and co-workers show that these designed zinc finger proteins can repress or activate expression of the endogenous genes *erbB-2* and *erbB-3* in living mouse, monkey, and human cells. *ErbB-2* is frequently overexpressed in human breast and ovarian cancers, and expression of *erbB-3* is suspected to be associated with cancer. In *erbB-2* and *erbB-3*, the 18-base-pair sites targeted by the synthetic proteins share 15 base pairs of common sequence, yet the transcription factors are specific enough to regulate one and not the other. The researchers also demonstrate that their transcription factors, and hence gene expression, can be regulated with the drug tetracycline.

Staff scientist Roy Pollock of Ariad Gene Therapeutics, Cambridge, Mass., who specializes in eukaryotic transcriptional regulation, comments that the sequence specificity of the transcription factors in the study is very impressive. He believes the Barbas study is the first to demonstrate both inhibition and activation of endogenous genes by designed transcription factors, rather than inhibition alone.

The findings suggest the possible use of designed transcription factors for gene therapy of cancer and other conditions. "Novartis has taken a license to our technology for the control of genes in plants," Barbas says, and Sangamo BioSciences, Richmond, Calif., "holds a license for the use of these proteins in human therapy."

Sangamo's president and chief executive officer, Edward O. Lanphier, points out that in unpublished studies the company has demonstrated the ability of engineered zinc finger DNA-binding proteins to control a number of endogenous genes. Sangamo has collaborative agreements with several other companies, including one with Baxter Healthcare to develop and commercialize zinc finger DNA-binding proteins that activate vascular endothelial growth factors (VEGFs) and VEGF receptors for treatment of vascular and cardiovascular diseases.

Several other research groups have developed or are developing related strategies for controlling gene expression. About six years ago, Klug and coworkers, including staff scientist Yen Choo, designed a zinc finger protein that bound specific DNA sequences in a leukemia gene and stopped uncontrolled cell



Barbas and coworkers created this model of a designed transcription factor with six zinc finger units (blue) bound to a DNA double helix (orange and red). They have designed transcription factors that can recognize a specific DNA sequence and repress or activate the gene containing that sequence.

growth in live mouse cells [*Nature*, 372, 642 (1994)]. This was the first report of a DNA-binding protein engineered *de novo* to inhibit gene expression.

The group of Carl O. Pabo, professor of biophysics and structural biology and a Howard Hughes Medical Institute investigator at Massachusetts Institute of Technology, has developed phage-display methods to identify zinc finger proteins that recognize specific sequences. Pabo and coworkers have also fused zinc finger motifs to other DNA-binding elements to create hybrid transcription factors with novel sequence specificities—a technology to which Ariad Gene Therapeutics has obtained an exclusive license. One such hybrid has been used successfully by Ariad researchers to bind to and activate synthetic therapeutic genes introduced into animals.

Ariad also is developing a technique in which a DNA-binding domain and an activation domain come together, in the presence of an orally available drug, to form a transcription factor. The idea is to eventually be able to use such transcription factors to turn synthetic genes on and off in patients for gene therapy applications.

And Dervan and coworkers have developed cell-permeable polyamides that fold into hairpin loops and undergo sequence-specific binding to base pairs in the minor groove of DNA, where they can play a regulatory role. Binding of these polyamides modulates gene expression by blocking other proteins, such as transcription factors, that would otherwise be able to bind to those sites. The researchers demonstrated three years ago that polyamides targeted at key sequences in the promoter region of a human immunodeficiency virus gene strongly inhibited viral replication [*Proc. Natl. Acad. Sci. USA*, 95, 12890 (1998)].

In unpublished work, Dervan's group and that of professor Mark S. Ptashne, head of the Gene Regulation Laboratory at Sloan-Kettering Institute, New York City, have created artificial transcription factors in which a small peptide activation domain is combined with a DNA-binding polyamide. The collaborative team finds that these synthetic transcription factors are capable of activating gene transcription *in vitro*. Dervan says he is particularly excited about these findings since previous polyamide constructs have been able to repress gene expression but not to activate it.

Dervan's synthetic polyamides can zero in on most any DNA sequence. The Barbas group's designed zinc finger pro-

teins can't do that yet. They bind up to 18-unit chunks of repeating GNN triplet sequences, where G is guanine and N is any of the other three types of DNA bases. Barbas and coworkers have reported the development of 16 distinct types of zinc finger domains, but 64 such domains would be required to bind any sequence at will.

Nevertheless, Barbas says, "it turns out that stretches of sequence that repeat the GNN motif six times are actually quite common in genes, and in every gene we've looked at we can find at least one of those sites. And we don't have to place the protein at very specific sites within the gene to regulate it. So we should be able to control most genes with our 16 zinc finger domains, and we believe that many of the 64 codon-recognition domains will become accessible in the next year or two." Barbas notes that in unpublished studies his group has prepared many of the additional domains required for the recognition of any given sequence.

Choo, Klug, and coworkers have already developed a new zinc finger display strategy that makes it possible to target more or less any 18-base-pair DNA sequence—not just repeating

GNN triplets. A scientific paper on the new strategy "is already written up and ready to go off," Choo notes, but some of the information has already appeared in patent applications filed by Gendaq Ltd., London, a company founded last year by Choo and Klug. At Gendaq, "we aim to apply customized transcription factors for gene regulation in functional genomics, agricultural biotechnology, and human therapy," Choo says.

With regard to potential applications of designed transcription factors, Barbas points out that gene therapy trials have been carried out in which patients with heart ailments have been injected with viruses containing VEGF genes. An alternative approach, he says, "would be simply to activate the endogenous VEGF gene found in every cell with a designed transcription factor." Furthermore, "you could imagine introducing a zinc finger protein that activates insulin and is under the control of a nontoxic, orally available pill," such as an aspirin-like derivative.

"We have encoded in our own genes the solutions to many diseases," Barbas says. "What awaits is just a way to switch on those critical genes."

Stu Borman

Enzyme's Activity Designed To Order

Researchers in England have modified a natural enzyme into a designed enzyme with an entirely new catalytic function. They accomplished this by cutting two sections out of the gene for the natural enzyme, inserting replacements containing mutations, shuffling the mutations, and screening expressed libraries of modified proteins for the desired activity—a process called directed evolution [*Nature*, 403, 617 (2000)].

Alan R. Fersht and coworkers at Cambridge Centre for Protein Engineering and Cambridge University Chemical Laboratory created the designed enzyme by substituting modified catalytic or binding units into an enzyme with an α/β -barrel structure—a common type of protein structural foundation or "scaffold."

In the study, Fersht and coworkers aimed to convert the activity of indole-3-glycerol phosphate synthase (IGPS) into that of phosphoribosylanthranilate isomerase (PRAI). The deck was loaded somewhat in that the two enzymes catalyze sequential *in vivo* reactions in the tryptophan biosynthetic pathway and PRAI's catalytic product is IGPS's sub-

strate. Nevertheless, the activities of the two enzymes are quite different, so it was no easy feat to get IGPS to mimic PRAI's catalytic function.

To do so, Fersht and coworkers snipped substrate-binding and catalytic loops out of the IGPS gene, inserted modified sequences in the gaps, shuffled the mutations, and screened the expressed enzyme variants for PRAI-like activity by testing them in PRAI-deficient bacteria that need tryptophan to grow. In the end, the researchers identified a modified IGPS with an activity and catalytic efficiency strikingly similar to those of PRAI.

The work has potential applicability to the creation of novel biocatalysts, and it advances scientists' understanding of the type of natural evolutionary mechanisms that enzymes might use to develop new functions and adapt to changing environments. The findings also suggest that the α/β -barrel scaffold—which may be found in as many as 10% of all soluble enzymes—could be generally useful for creating biocatalysts with a range of novel activities.

"Alan was trying to demonstrate that

The opinion in support of the decision being entered today was not written
for publication and is not binding precedent of the Board.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

**Ex parte GEORGE NORBERT COX III, CASEY CHRISTOPHER CASE,
STEPHEN P. EISENBERG, ERIC EDWARD JARVIS and
SHARON KAYE SPRATT**

Appeal No. 2006-1270
Application No. 10/222,614

MAILED

SEP 27 2006

U.S. PATENT AND TRADEMARK OFFICE
BOARD OF PATENT APPEALS
AND INTERFERENCES

ON BRIEF

Before GRIMES, GREEN, and LEBOVITZ, Administrative Patent Judges.

GREEN, Administrative Patent Judge.

DECISION ON APPEAL

This is a decision on appeal under 35 U.S.C. § 134 from the examiner's final rejection of claims 114, 116 and 119-124, all of the pending claims. Claims 114 and 120-124 are representative of the claims on appeal, and read as follows:

114. A cell comprising first and second engineered zinc finger proteins, where each of the zinc finger proteins further comprises an endonuclease or functional fragment thereof, and further wherein:
(a) the first protein binds to a first target site; and
(b) the second protein binds to a second target site.

RECEIVED

OCT 02 2006

SANGAMO BIO SCIENCES, INC.

120. The cell of claim 114, wherein the cell is an animal cell.
121. The cell of claim 120, wherein the cell is a mammalian cell.
122. The cell of claim 121, wherein the cell is a human cell.
123. The cell of claim 122, wherein the cell is a stem cell.
124. The cell of claim 123, wherein the cell is a hematopoietic stem cell.

Claims 123 and 124 stand rejected under 35 U.S.C. § 112, first paragraph, as failing to comply with the enablement requirement. In addition, Claims 114, 116 and 119-122 stand rejected under 35 U.S.C. § 103(a) as being obvious over the combination of Brenneman¹ and Chandrasegaran.² After careful review of the record and consideration of the issues before us, we reverse both rejections.

DISCUSSION

Claims 123 and 124 stand rejected under 35 U.S.C. § 112, first paragraph, "as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention." Examiner's Answer, page 3.

¹ Brenneman et al. (Brenneman), "Stimulation of intrachromosomal homologous recombination in human cells by electroporation with site-specific endonucleases," *PNAS*, Vol. 93, pp. 3608-12 (1996).

² Chandrasegaran, U.S. Pat. No. 5,792,640, issued August 11, 1998.

"[A] specification disclosure which contains a teaching of the manner and process of making and using the invention in terms which correspond in scope to those used in describing and defining the subject matter sought to be patented must be taken as in compliance with the enabling requirement of the first paragraph of § 112 unless there is reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support." In re Marzocchi, 439 F.2d 220, 223, 169 USPQ 367, 369 (CCPA 1971) (emphasis in original). "[It] is incumbent upon the Patent Office, whenever a rejection on this basis is made, to explain why it doubts the truth or accuracy of any statement in a supporting disclosure and to back up assertions of its own with acceptable evidence or reasoning which is inconsistent with the contested statement." Id. at 224, 169 USPQ at 370. Here, the examiner has not provided "acceptable evidence or reasoning which is inconsistent" with the specification, and therefore has not met the initial burden of showing nonenablement.

The examiner engages in a Wands analysis, see In re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400, 1403 (Fed. Cir. 1988) (noting that facts that should be considered in determining whether a specification is enabling include: (1) the quantity of experimentation necessary to practice the invention, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claims) to come to the conclusion that it would require

an undue amount of experimentation to make and/or use the claimed cells, wherein the cells are stem cells. See Examiner's Answer, pages 3-6.

The examiner notes that there are no working examples, and that the "specification provides guidance to use zinc finger-endonuclease fusion proteins as gene regulators . . . and as transcription repressors . . . without further elaboration as to how such functions can be achieved." Id. at 4. Those are issues that would seem to apply to the use of any cell type, not just stem cells, and the examiner does not explain how those issues provide more of an impediment to the use of stem cells as opposed to other cell types.

With respect to stem cells, the examiner focuses on the use of the zinc finger-endonuclease constructs in homologous recombination in stem cells. See id. at 4. The examiner cites Hatada³ for the proposition "that hematopoietic stem cells have not been shown to perform homologous recombination." Id. Hanson⁴ is cited for demonstrating that "hematopoietic stem cells are difficult to purify and manipulate." Id. at 4-5. Finally, Zwaka⁵ is cited for the discussion that "human embryonic stem cells are more difficult to manipulate than prior art mouse embryonic stem cells," and that it does not appear that homologous

³ Hatada et al. (Hatada), "Gene correction in hematopoietic progenitor cells by homologous recombination," PNAS, Vol. 97, No. 25, pp. 13807-811 (2000).

⁴ Hanson et al. (Hanson), "Enhanced green fluorescent protein targeted to the Sca-1 (*Ly-6A*) locus in transgenic mice results in efficient marking in hematopoietic stem cells *in vivo*," Experimental Hematology, Vol. 31, pp. 159-67 (2003).

⁵ Zwaka et al. (Zwaka), "Homologous recombination in human embryonic stem cells," Nature Biotechnology, Vol. 21, pp. 319-21 (2003).

recombination was known in any other human stem cell at the time of its publication. Id. at 5.

The examiner also states that

[t]he prior art does not show human stem cells with zinc finger-endonuclease fusion proteins, the prior art therefore does not predict whether such cells could be made or used. Hatada [] shows that stem cells are difficult to isolate, Hanson [] shows that hematopoietic stem cells are difficult to purify and manipulate, and Zwaka [] show[s] that human embryonic stem cells are difficult to manipulate.

Id.

As noted by appellants, see Appeal Brief,⁶ page 5, the Hatada and Zwaka references relate to homologous recombination, and the claims are not limited to cells containing two engineered zinc finger proteins that also comprise an endonuclease, that are required to undergo homologous recombination. Cells comprising a single zinc finger-nuclease fusion protein, appellants assert citing Chandrasegaran, "were known to be useful in mutagenesis, targeted cleavage, gene expression, detection of conformational changes in nucleic acid and targeted recombination." Reply Brief, page 4. Thus, Hatada and Zwaka, as they relate to the frequency of homologous recombination in human stem cells, are not relevant to the issue of whether claims 123 and 124 are enabled by the specification. Moreover, while Zwaka teaches that electroporation protocols that have been developed for mouse embryonic stem cells do not achieve the same results in human embryonic stem cells, see id., abstract, the reference teaches

⁶ All references to the "Appeal Brief" are to the Appeal Brief dated July 11, 2005.

further that “[f]or human embryonic stem cells, the best chemical reagents yield stable (drug-selectable) transfectants at rates about 10^{-5} ,” id. at page 319, first column, second full paragraph. Thus, Zwaka teaches that human embryonic stem cells may be transfected through the use of chemical reagents.

Moreover, while Hanson teaches that “[e]xperimental manipulation of hematopoietic stem cells is challenging . . . [as] [t]hey are difficult to purify, propagate ex vivo, assay, and transduce,” id. at 159, second column, Hanson also teaches that enhanced green fluorescent protein was integrated into the Sca-1 (glycosyl phosphatidyl-anchored protein) locus by homologous recombination in mouse embryonic stem cells, see id., abstract. Thus, Hanson demonstrates while it may be difficult to manipulate hematopoietic stem cells, it is possible to do so. See, e.g., Johns Hopkins University v. CellPro, Inc., 152 F.3d 1342, 136-61, 47 USPQ2d 1705, 1719 (Fed. Cir. 1998) (“The test [for undue experimentation] is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed to enable the determination of how to practice a desired embodiment of the claimed invention.” (insert in original)).

Finally, the fact that the prior art does not show human stem cells with zinc finger-endonuclease fusion proteins is not the correct standard to measure enablement, for if it were, any novel and/or non-obvious invention would be, by definition, non-enabled.

Therefore, as the examiner has failed to set forth a prima facie case of unpatentability under 35 U.S.C. § 112, first paragraph, we are compelled to reverse the rejection.

Claims 114, 116 and 119-122 stand rejected under 35 U.S.C. § 103(a) as being obvious over the combination of Brenneman and Chandrasegaran.

The panel would first like to note that the rejection is premised on an incorrect claim construction.

Claim 114 is drawn to “[a] cell comprising first and second engineered zinc finger proteins, where each of the zinc finger proteins further comprises an endonuclease or functional fragment thereof, and further wherein: (a) the first protein binds to a first target site; and (b) the second protein binds to a second target site.”

According to the examiner, “[t]he phrase ‘A cell comprising first and second engineered zinc finger proteins’ . . . is interpreted to include cells with two identical zinc finger proteins. The phrase ‘(a) the first protein binds to a first target site; and (b) the second protein binds to a second target site’ . . . is interpreted to include two identical zinc finger target sites.” Examiner’s Answer, page 6.

The problem with the examiner’s construction is that it is reading the limitations “first and second engineered zinc finger protein” and “a first target site” and “a second target site” out of the claims. We construe “first and second engineered zinc finger protein” as two distinct and different zinc finger proteins,

and construe "a first target site" and "a second target site" as two distinct and different target sites. Now that the claims have been construed, we now turn our attention to review of the obviousness rejection.

Brenneman is cited for disclosing "that the efficiency of homologous recombination in a human cell can be increased by digestion by an endonuclease at the site of homology." Examiner's Answer, pages 6-7. Brenneman specifically teaches that *Xba* I endonuclease, as well as the rare-cutting yeast endonuclease *PI-Sce* I increased the frequency of recombination, whereas restriction enzymes that cut outside of the repeated regions or between them "produced no change in recombination frequency." Brenneman, abstract. The examiner notes that "Brenneman [] does not show use of a chimeric nuclease that comprises a zinc finger protein." Examiner's Answer, page 7.

Chandrasegaran is cited for disclosing "bacterial cells transformed with a fusion protein of a three-zinc finger DNA binding domain linked to a catalytic nuclease domain of Fok I." Id. Chandrasegaran is also cited for teaching that each finger of the zinc finger protein binds to three nucleotides of a polynucleotide, and that zinc finger proteins may be designed to bind a series of triplet nucleotides of choice. See id.

The examiner concludes:

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the endonuclease used by Brenneman [] by use of the chimeric zinc finger-Fok I endonuclease of Chandrasegaran because use of the endonuclease of Chandrasegaran allows for cleavage at other repeated sites of choice and would thereby increase the frequency

of homologous recombination at the sites of choice. Recombination at repeated sites of choice would further enable generation of desired recombination products and allow for further study of the process of homologous recombination as exemplified by the experiments of Brenneman. Brenneman [] shows that cleavage at two repeated sites in a target sequence for homologous recombination increases the rate of homologous recombination, and further shows a large repeated sequence as a substrate for homologous recombination. Chandrasegaran shows that a chimeric zinc finger-Fok-I endonuclease may be designed to specifically bind to a sequence of 9 base pairs and that it is possible to design the zinc finger portion to bind to a target sequence of choice. It is further obvious from consideration of the effect of expression of endonucleases on cell viability in both Brenneman [] and Chandrasegaran that the lethality of the endonuclease used by both groups is due to cleavage at multiple sites in the host cell chromosome, especially in view of the observation by Brenneman [] that the rare-cutter PI-Sce I is not lethal to the host cell. Therefore both the cells of Brenneman [] and Chandrasegaran show cells with multiple target sites for endonucleolytic cleavage.

Id. at 8.

Appellants argue that “[N]owhere do the cited references, alone or in combination, teach or suggest the cells including two chimeric zinc fingers as claimed.” Appeal Brief, page 9. Moreover, appellants assert that “the Examiner’s interpretation that the claims ‘include cells with two identical zinc finger proteins’ and ‘two identical zinc finger target sites’ is not correct . . . [as] throughout prosecution, Appellants have repeatedly characterized the claimed subject matter as relating to ‘cells comprising two zinc finger proteins of different sequence.’” Reply Brief, page 8. We agree, and the rejection must be reversed.

As we have stated above, the examiner’s rejection is based on an erroneous claim construction. And as the examiner has not pointed to how the

Sean Brennan
Sangamo BioSciences, Inc.
501 Canal Blvd.
Suite A100
Richmond CA 94804



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/984,304 ✓	11/09/2004 ✓	George Norbert Cox III	8325-0002.04	5124
20855	7590	09/26/2008	EXAMINER	
ROBINS & PASTERNAK 1731 EMBARCADERO ROAD SUITE 230 PALO ALTO, CA 94303			BRUSCA, JOHN S	
		ART UNIT	PAPER NUMBER	
		1631		
		MAIL DATE		DELIVERY MODE
		09/26/2008		PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

RECEIVED

SEP 29 2008

ROBINS & PASTERNAK LLP

DOCKETED *D&P*

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte GEORGE NORBERT COX III,
CASEY CHRISTOPHER CASE, STEPHEN P. EISENBERG,
ERIC EDWARD JARVIS, and SHARON KAYE SPRATT

Appeal 2008-4125
Application 10/984,304
Technology Center 1600

Decided: September 26, 2008

Before TONI R. SCHEINER, DONALD E. ADAMS, and
FRANCISCO C. PRATS, *Administrative Patent Judges*.

PRATS, *Administrative Patent Judge*.

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134 involving claims to a method for modulating expression of a viral gene in a virally infected cell. The Examiner has rejected the claims as being indefinite, obvious, and for obviousness-type double patenting. We have jurisdiction under 35 U.S.C. § 6(b).

Appeal 2008-4125
Application 10/984,304

We affirm the obviousness and obviousness-type double patenting rejections, but remand the case to the Examiner to reconsider the indefiniteness rejection.

STATEMENT OF THE CASE

Claims 70-72 are pending and on appeal (App. Br. 2). Claim 70 is representative and reads as follows:

70. A method for modulating expression of a viral gene in a virally infected cell, wherein the method comprises:
- (a) expressing a zinc finger protein in the cell, wherein the zinc finger protein has been engineered to bind to a target site in a viral gene; and
 - (b) maintaining the cell under conditions in which the engineered zinc finger protein binds to a target site in the viral gene.

The Examiner applies the following documents in rejecting the claims:

Saiga et al. US 6,090,783 Jul. 18, 2000

E. H. Nasser et al., *Antiviral Activity of Influenza Virus M1 Zinc Finger Peptides*, 70 JOURNAL OF VIROLOGY 8639-8644 (December 1996).

Appeal 2008-4125
Application 10/984,304

The following rejections¹ are before us for review:

Claims 70-72 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention (Ans. 4).

Claims 70-72 stand rejected under 35 U.S.C. § 103(a) as being obvious in view of Saiga and Nasser (Ans. 4-6).

Claims 70-72 stand provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 5, 7, and 12 of copending Application No. 11/148,794 (Ans. 6).

OBVIOUSNESS

ISSUE

The Examiner cites Saiga as disclosing “a human zinc finger protein termed TRP-1 that comprises a KRAB transcriptional repression domain in columns 20-24. Saiga et al. shows that control of gene expression in human T-cell leukemia virus type 1 (HTLV-1) is controlled by the promoter in the LTR region of HTLV-1 in column 2” (Ans. 4-5). The Examiner cites Example 8 of Saiga as demonstrating that expressing the zinc finger protein

¹ Appellants filed an after-final amendment on January 8, 2008, which the Examiner declined to enter (App. Br. 3-4). Appellants urge that the Examiner’s basis for denying entry was improper, and that the Board should consider the merits of the Examiner’s basis for non-entry of the amendment (*see id.* at 4-7). However, because the amendment was not entered, the claim proposed by Appellants was not subject to any ground of rejection, and is therefore not eligible for appeal. *See* 35 U.S.C. § 134(a) (claims must be twice rejected before appeal). Moreover, refusal to enter an amendment is a petitionable matter not decided by the Board. *See* MPEP § 1002.02(c); 37 C.F.R. §§ 1.127, 1.181. We therefore will not discuss the merits of the non-entered amendment.

TRP-1 in HeLa cells via a transfected expression vector inhibits transcription of cotransfected HTLV-1 genes (*id.* at 5).

The Examiner concedes that Saiga "does not show use of a zinc finger protein to repress expression of a viral gene in an infected cell," and cites Nasser to meet that limitation (*id.*). Specifically, the Examiner cites Nasser as disclosing "a zinc finger peptide termed peptide 6 that is derived from the influenza virus M1 protein. Nasser et al. shows on page 8639 that peptide 6 was chemically synthesized, and comprises amino acids 148-166 of the wild type M1 protein" (*id.*).

The Examiner finds that Nasser discloses that peptide 6 represses infection of cultured cells by influenza virus, is 1000-fold more effective than the wild-type M1 protein in inhibiting influenza transcriptase, and that "[p]eptide 6 may provide a new approach to the design of antiviral agents effective against influenza virus and possibly other viruses" (*id.* at 5-6 (quoting Nasser 8644)). Based on these teachings, the Examiner concludes that one of ordinary skill in the art would have considered it obvious to modify Saiga's method "to repress transcription of HTLV-1 viral genomes in infected cells because Nasser et al. shows that zinc finger proteins are capable of blocking viral infections and both Saiga et al. and Nasser et al. provide guidance to use zinc finger proteins as antiviral agents" (*id.* at 6).

Appellants contend that "the term 'engineered to bind to a target site' . . . clearly refers to a non-naturally occurring zinc finger protein that has been altered (e.g., designed or selected) to bind to a particular target site" (App. Br. 10). Therefore, Appellants urge, "the pending claims are directed to [a] method of modulating viral infection using a non-naturally occurring zinc

finger protein including zinc finger DNA recognition domains that have been engineered to bind to a target site in a viral gene" (*id.*).

In contrast, Appellants argue, Saiga and Nasser "are completely silent" regarding zinc finger proteins "that have been engineered to bind to a target site as recited in the pending claims and, indeed, teach away from such proteins, clearly teaching that their proteins comprise non-engineered (naturally-occurring) DNA-binding domains so as to ensure the desired function by binding to the cognate target sites" (*id.*). Moreover, Appellants argue, the Examiner's conclusion of obviousness is erroneous because both Saiga and Nasser use naturally occurring proteins to inhibit viral activity; therefore, modifying either reference to obtain the claimed non-naturally-occurring engineered proteins would destroy the intended function of the prior art proteins (*id.* at 10-12).

Appellants do not argue any of the claims separately. We select claim 70 as representative of the rejected claims. 37 C.F.R. § 41.37(c)(1)(vii).

The issue with respect to this rejection, then, is whether the Examiner has established a *prima facie* case that one of ordinary skill in the art would have considered it obvious, in view of Saiga and Nasser, to modulate the expression of a viral gene in a virally infected cell by expressing in the cell a zinc finger protein that has been engineered to bind to a target site in a viral gene, as recited in claim 70.

FINDINGS OF FACT ("FF")

1. Saiga discloses "[a] protein (TRP-1) which binds to a transcriptional repressive region existing in the U5 region of human T-cell leukemia virus type I gene LTR . . . [and] includes a domain common to Kruppel-type

transcriptional repressive factors and five Kruppel-type zinc finger domains” (Saiga, col. 4, ll. 43-48).

2. Saiga discloses that the TRP-1 protein “specifically binds to U5RE existing in the U5 region of human T-cell leukemia virus type I gene LTR” (Saiga, col. 4, ll. 31-32).
3. Saiga discloses that the TRP-1 protein and its gene were obtained by expressing a cDNA library of the human acute lymphocytic leukemia cell line Molt-4, determining which proteins bound U5RE sequences, and obtaining the appropriate clone from the library (Saiga, col. 20, l. 10 through col. 22, l. 64 (Example 5)).
4. In order to test whether expression of TRP-1 inhibited transcription of HTLV-1 genes, Saiga discloses, in Example 8, introducing nucleic acid constructs into HeLa cells as follows:

An expression vector pEF-HA-TRP-1 obtained by engineering an EF-BOS vector so that a HA-TRP-1 fusion protein having the influenza HA [(hemagglutinin)] tag at the N-terminus of TRP-1 would be expressed, and a reporter plasmid TK-CAT in which HSV TK (a minimum promoter region) was linked upstream of the CAT (chloramphenicol acetyl transferase) gene or TK-3xU5RE-CAT in which three U5REs were inserted between the TK and CAT genes, were simultaneously introduced into HeLa cells....

(Saiga, col. 23, l. 67, through col. 24, l. 8.)

5. Saiga discloses that “[t]he TK-3xU5RE-CAT includes U5REs, which are binding sequences for TRP-1, whereas TK-CAT includes no binding sequences” (Saiga, col. 24, ll. 9-11). Thus, in the experiment described in Example 8, “a pair consisting of pEF-HA-TRP-1 and TK-3xU5RE-CAT or a pair consisting of pEF-HA-TRP-1 and TK-CAT was introduced into HeLa

cells so as to analyze whether or not TRP-1 functions via U5RE" (*id.* at col. 24, ll. 37-41 (emphasis added)).

6. Saiga discloses that upon culturing the cells it was found that "the CAT activity by TK-3xU5RE-CAT was reduced by 35% in a concentration-dependent manner based on the concentration of the pEF-HA-TRP-1 plasmid, whereas no effect was observed for TK-CAT. Thus, it was indicated that TRP-1 has transcription repression activity via USRE" (Saiga, col. 24, ll. 41-46).

7. Saiga discloses that the TRP-1 protein disclosed therein can be modified in a number of ways and still function in accordance with the disclosure:

A "sequence similar to" an amino acid sequence or a DNA sequence is not limited to any particular sequence, but is defined as such a sequence modified with substitutions, insertions, deletions, and the like known to those skilled in the art so that the function or activity of its encoded protein is substantially at the same level. Or, as long as the function or activity of the protein is substantially at the same level, it may contain chemical or biochemical modifications, or non-natural or derivatized amino acids or bases. For example, the above-mentioned TRP-1 protein preferably has similarity of about 50% or more, or homology of about 35% or more with the natural type. More preferably, the TRP-1 protein has similarity of about 70% or more, or homology of about 50% or more with the natural type. Still more preferably, the TRP-1 protein has similarity of about 80% or more; or homology of about 65% or more. Herein, "similarity" is defined as the rate (%) of identical amino acids within a similar sequence with respect to a reference sequence, where the amino acids are divided into the following five groups A to E and amino acids within each group are considered as identical; group A: Ala, Ser, Thr, Pro, and Gly; group B: Asn, Asp, Glu, and Gin; group C: His, Arg, and Lys; group D: Met, Leu, Ile, and Val; and group E: Phe,

Tyr, and Trp. The "homology" of an amino acid sequence is defined as the rate (%) of identical amino acids within a similar sequence with respect to a reference sequence, where only completely identical amino acids are considered as identical. *Furthermore, the "homology" of a DNA sequence is not limited to any particular sequence, but is defined as such a sequence modified with substitutions, insertions, deletions, and the like, known to those skilled in the art, especially so that the function of the DNA sequence, e.g., gene expression repressing function for HTLV-I, is substantially at the same level.*

(Saiga, col. 7, ll. 19-52 (emphases added).)

8. Nasser discloses that the influenza virus protein M1 "has been shown to inhibit influenza virus transcriptase" (Nasser 8639).
9. Nasser discloses the chemical synthesis of a peptide, termed peptide 6, composed of amino acid residues 148 to 166 of the M1 protein (Nasser 8639).
10. Nasser discloses that peptide 6 "represents a Zn²⁺ finger which includes a 7-residue 'loop' and a 4-residue 'tail' in addition to the 4 residues on either side of the loop involved in coordination of Zn²⁺" (Nasser 8639).
11. Nasser discloses that "[w]hen the peptide was introduced into tissue culture 5 min after viral challenge with A/PR/8/34, antiviral activity was seen at levels as low as 0.1 nM; on a molar basis, the peptide was shown to be 1,000- to 2,500-fold more effective than ribavirin or amantadine," and that "[a]ntiviral activity was seen with addition of the peptide up to 1 h after viral infection" (Nasser 8639).
12. Nasser discloses that peptide 6 "was 1,000-fold more effective on a molar basis in transcriptase inhibition than was M1" (Nasser 8639).
13. Nasser discloses that "[i]n vivo studies have been performed with peptide 6, using a mouse influenza model; when administered intranasally,

peptide 6 was found to be as active as ribavirin against A/PR/8/34 (H1N1) and more active than ribavirin against A/Victoria/3/75 (H3N2)" (Nasser 8643-44).

14. Nasser discloses that peptide 6 "may provide a new approach to the design of antiviral agents effective against influenza virus and possibly other viruses" (Nasser 8644).

PRINCIPLES OF LAW

Recently addressing the question of obviousness, the Supreme Court reaffirmed that under the controlling inquiry, "the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved." *KSR Int'l Co. v. Teleflex Inc.*, 127 S. Ct. 1727, 1734 (2007) (quoting *Graham v. John Deere Co. of Kansas City*, 383 U.S. 1, 17-18 (1966)).

In proceedings before the Patent and Trademark Office, the Examiner bears the burden of establishing a *prima facie* case of obviousness based upon the prior art. "[The Examiner] can satisfy this burden only by showing some objective teaching in the prior art or that knowledge generally available to one of ordinary skill in the art would lead that individual to combine the relevant teachings of the references."

In re Fritch, 972 F.2d 1260, 1265 (Fed. Cir. 1992) (citations omitted, bracketed material in original). Thus, as the Supreme Court pointed out in *KSR*, "a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art." *KSR*, 127 S. Ct. at 1741.

While holding that some rationale must be supplied for a conclusion of obviousness, the Court nonetheless rejected a "rigid approach" to the

obviousness question, and instead emphasized that “[t]hroughout this Court's engagement with the question of obviousness, our cases have set forth an expansive and flexible approach” *Id.* at 1739. The Court also rejected the use of “rigid and mandatory formulas” as being “incompatible with our precedents.” *Id.* at 1741; *see also* 1742-43 (“Rigid preventative rules that deny factfinders recourse to common sense, however, are neither necessary under our case law nor consistent with it.”).

The Court thus reasoned that the analysis under 35 U.S.C. § 103 “need not seek out precise teachings directed to the specific subject matter of the challenged claim, for a court can take account of the inferences and creative steps that a person of ordinary skill in the art would employ.” *Id.* at 1741. The Court further advised that “[a] person of ordinary skill is . . . a person of ordinary creativity, not an automaton.” *Id.* at 1742.

Regarding hindsight reasoning, the Court stated that “[a] factfinder should be aware, of course, of the distortion caused by hindsight bias and must be cautious of arguments reliant upon *ex post* reasoning. Rigid preventative rules that deny factfinders recourse to common sense, however, are neither necessary under our case law nor consistent with it.” *Id.* at 1742-1743 (citations omitted).

ANALYSIS

Appellants' arguments do not persuade us that the Examiner has failed to establish a *prima facie* case of obviousness with respect to claim 70. Rather, we agree with the Examiner that a person of ordinary skill in the art viewing Saiga and Nasser would have considered it obvious to modulate the expression of a viral gene in a virally infected cell by expressing in the cell a

zinc finger protein that has been engineered to bind to a target site in a viral gene.

Specifically, Saiga would have advised one of ordinary skill in the art that the TRP-1 protein was capable of specifically binding to an HTLV-1 repressor region (FF 1-3), and that, when the gene encoding the TRP-1 was expressed in a cell, the TRP-1 protein's binding capacity rendered it capable of inhibiting expression of viral genes (*see* FF 4-6). While Saiga does not appear to disclose expressing the TRP-1 protein in a virally infected cell, Nasser would have advised one of ordinary skill in the art that zinc finger proteins were capable of exerting antiviral effects in virally infected cells at dosages comparable to other antiviral agents (FF 10-14). Given these teachings, we agree with the Examiner that one of ordinary skill in the art would have reasonably inferred that it would be desirable to express the TRP-1 protein in virally infected cells so as to repress the viral genes.

With respect to the disputed limitation, claim 70 recites that "the zinc finger protein has been engineered to bind to a target site in a viral gene." Appellants argue that this limitation "clearly refers to a non-naturally occurring zinc finger protein that has been altered (e.g., designed or selected) to bind to a particular target site" (App. Br. 10). Because both references use a naturally-occurring zinc finger protein, Appellants argue, neither reference teaches or suggests this limitation (*id.*)

The Examiner responds that the limitation "engineered to bind to a target site in the viral gene" is essentially a product-by-process limitation on the protein, and therefore encompasses even wild-type (i.e. naturally-occurring) proteins because "[a]ny wild type sequence of a zinc finger protein could be arrived at by a process of modification of a related, but

different, zinc finger protein" (Ans. 8). Thus, the Examiner argues, "[t]he zinc finger protein shown in Saiga et al. meets all the functional limitations of the claimed subject matter, and therefore Saiga et al. shows the zinc finger used in the claimed processes" (*id.*).

The Examiner further contends that, even if claim 70 is limited to the use of non-naturally-occurring zinc finger proteins, the cited references meet that limitation because the protein expressed by Saiga is a TRP-1/HA fusion protein, and because Nasser uses a chemically synthesized peptide that is obtained from, and therefore different than, the wild-type influenza virus M1 protein (Ans. 8-9 (citing Saiga, columns 23-24 (*see FF 4*) and Nasser 8639 (*see FF 9-13*)). Appellants respond:

[T]he evidence of record establishes that the engineering to bind to a target site is accomplished by modification (design or selection) of the amino acid sequence of the recognition helix, not fusion to a heterologous domain. Accordingly, the claimed zinc finger proteins are themselves non-naturally occurring. By contrast, a zinc finger protein with an unaltered (non-engineered) DNA-binding binding domain is still a naturally occurring zinc finger protein even when incorporated into a (non-naturally occurring) fusion protein.

(Reply Br. 8.)

Even accepting the narrow definition advanced by Appellants, we are not persuaded that the cited references fail to suggest expressing a "zinc finger protein . . . engineered to bind to a target site in the viral gene" to modulate viral gene expression in a virally infected cell. Specifically, Saiga explicitly discloses that the HTLV-1-repressing TRP-1 protein can be modified in a number of ways and still function in accordance with the disclosure (FF 7).

Thus, Saiga discloses that the TRP-1 protein is not limited to any particular naturally-occurring sequence, but instead can be “modified with substitutions, insertions, deletions, and the like known to those skilled in the art so that the function or activity of its encoded protein is substantially at the same level” (Saiga, col. 7, ll. 21-24 (FF 7)). Saiga further discloses that the DNA sequence encoding the TRP-1 protein is not “limited to any particular sequence, but is defined as such a sequence modified with substitutions, insertions, deletions, and the like, known to those skilled in the art, especially so that the function of the DNA sequence, e.g., gene expression repressing function for HTLV-I, is substantially at the same level” (Saiga, col. 7, ll. 47-52 (FF 7)).

Thus, even if one accepts claim 70’s recitation “zinc finger protein . . . engineered to bind to a target site in the viral gene” to mean “a non-naturally occurring zinc finger protein that has been altered (e.g., designed or selected) to bind to a particular target site” (App. Br. 10), Saiga explicitly discloses that its zinc finger protein can be altered with non-naturally occurring substitutions and deletions in the amino acid sequence, and still have its DNA-binding, gene-repressing functionality. We therefore do not agree that the cited references fail to provide any teaching or suggestion of a zinc finger protein that meets the definition advanced by Appellants.

While it is noted that Saiga contemplates non-naturally occurring modifications that result in substantially the same level of activity as the wild-type protein, claim 70 does not contain any limitation requiring the “engineering” to result in a higher activity than that present in the natural protein. Moreover, because Saiga explicitly contemplates non-naturally occurring modifications to its zinc finger protein, we are not persuaded that

one of ordinary skill in the art would find that making such modifications would destroy the protein's intended function.

Thus, we agree with the Examiner that one of ordinary skill in the art would have considered claim 70 *prima facie* obvious in view of Saiga and Nasser, even when the recitation "zinc finger protein . . . engineered to bind to a target site in the viral gene" is interpreted to mean "a non-naturally occurring zinc finger protein that has been altered (*e.g.*, designed or selected) to bind to a particular target site" (App. Br. 10), as urged by Appellants. We therefore affirm the Examiner's obviousness rejection of claim 70. Claims 71 and 72 fall with claim 70 because they were not argued separately. 37 C.F.R. § 41.37(c)(1)(vii).

OBVIOUSNESS-TYPE DOUBLE PATENTING

Claims 70-72 stand provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 5, 7, and 12 of copending Application No. 11/148,794 (Ans. 6).

The Examiner contends that while the appealed claims and conflicting claims in the copending application are not identical, "they are not patentably distinct from each other because the copending claims of Application No. 11/148,794 are drawn to species of the instant method of modulation of viral gene expression with regards to the number and randomization of the zinc finger domains in the protein" (*id.*). Appellants present no substantive argument regarding this rejection, and instead "request that the obviousness-type double patenting rejection be held in abeyance pending indication of allowable claims in either application" (App. Br. 12).

In the absence of any argument that the Examiner's rejection is erroneous, we affirm the Examiner's provisional obviousness-type double patenting rejection of claims 70-72 over claims 5, 7, and 12 of copending Application No. 11/148,794.

INDEFINITENESS

Claims 70-72 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention (Ans. 4).

The Examiner states that "the phrase 'wherein the zinc finger protein has been engineered to bind to a target site in a viral gene'" is indefinite because the term "'engineered' is not defined in the specification and it is not clear how it limits the structure of the recited zinc finger protein" (*id.*).

Appellants respond that reference to the Specification shows that "the recitation 'wherein the zinc finger protein has been engineered to bind to a target site in a viral gene' . . . clearly refers to zinc finger proteins which have been altered in the recognition region helix by design or selection to bind to a selected target site" (App. Br. 8 (citing Spec. 10:17-21; Spec. 21: 10-20; Spec. 24:23-33; Spec. 74:13-16)).

We find this rejection unripe for appeal, and therefore remand the case to the Examiner to reconsider the merits of the indefiniteness rejection in light of the following discussion.

At the outset, we understand that before any decision on the merits of an appeal can be undertaken, the interpretation of the claims must be set. See *In re Steele*, 305 F.2d 859, 862 (CCPA 1962) ("[S]peculation as to the meaning of the terms employed and assumptions as to the scope of such claims" is legal error.); see also *In re Geerdes*, 491 F.2d 1260, 1262 (CCPA

1974) (“Before considering the rejections under 35 U.S.C. §[] 103 . . . we must first decide [what] the claims include within their scope.”).

However, regarding the appealed rejection under § 103, as discussed above, the cited references would have rendered the claimed subject matter obvious to a person of ordinary skill in the art regardless of whether we adopt the Examiner’s broad definition of “engineered,” or Appellants more limited definition.

Turning to the substance of the Examiner’ indefiniteness rejection, we first note that this application is a continuation of U.S. Patent Application Serial No. 09/897,844, filed July 2, 2001, which issued as U.S. Patent No. 6,979,539 B2, which is in turn a continuation of U.S. Patent Application Serial No. 09/229,037, filed January 12, 1999, which issued as U.S. Patent No. 6,534,261 B1. The claims of each of those issued patents contain a number of recitations regarding “engineered” zinc finger proteins.

Thus, despite having Specifications identical to that of the instant case, the Examiner has, in the instant case, concluded that the recitation “engineered” is indefinite -- a conclusion that appears to be directly inconsistent with the Examiner’s allowance of claims containing that term in the two parent cases. It is therefore unclear whether the Examiner has interpreted the language at issue in a consistent manner throughout the prosecution of these cases.

We therefore remand the case to the Examiner to ensure that the claims at issue herein are being interpreted in a manner consistent with the interpretation applied in the parent cases.

When the Examiner takes the case up for action the Examiner should take a step back, and review the prosecution histories in the parent cases to

ensure that the interpretation applied to the instant claims is consistent with the interpretation applied to the patented parent cases. The Examiner should also ensure that the position taken is consistent with the position one of ordinary skill in the art would take.

For example, the Examiner appears to have determined that one of ordinary skill in the art would consider “the term ‘engineering’ to describe a wide range of processes, including expression of a naturally occurring protein by recombinant DNA methods, and modifications of any portion of a gene or its expressed protein” (Ans. 7). Thus, despite having provided a definition for the term “engineering,” the Examiner nonetheless concludes that the term is indefinite. The Examiner is reminded that “breadth is not to be equated with indefiniteness.” *In re Miller*, 441 F.2d 689, 693 (CCPA 1971).

Also, even if the Examiner concedes that the Specification somehow limits the scope of the claims in the manner advanced by Appellants, the Examiner should consider whether, and why, that interpretation is indefinite.

If after reconsidering the rejection in light of the above discussion the Examiner should conclude that the rejection must be maintained, the Examiner should include in the rejection an explicit statement explaining why the indefiniteness rejection is consistent with the position taken in the parent cases.

SUMMARY

We affirm the Examiner’s rejection of claims 70-72 under 35 U.S.C. § 103(a) as obvious in view of Saiga and Nasser.

Appeal 2008-4125
Application 10/984,304

We affirm the Examiner's provisional obviousness-type double patenting rejection of claims 70-72 over 5, 7, and 12 of copending Application No. 11/148,794.

However, because it is not clear that the Examiner has interpreted the claims consistently with the prosecution in the parent cases, we remand the case to the Examiner to reconsider the indefiniteness rejection under 35 U.S.C. § 112, second paragraph, in accordance with the discussion set forth herein.

Because we find that the case must be remanded to the Examiner, we hold the finality of our affirmances of the obviousness and obviousness-type double patenting rejections in abeyance until the proceedings on remand before the Examiner are concluded. 37 C.F.R. § 41.50(e).

AFFIRMED, REMANDED

cdc

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO CA 94303



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/984,304	11/09/2004	George Norbert Cox III	8325-0002.04	5124
20855	7590	03/26/2009		
ROBINS & PASTERNAK 1731 EMBARCADERO ROAD SUITE 230 PALO ALTO, CA 94303			EXAMINER BRUSCA, JOHN S	
			ART UNIT 1631	PAPER NUMBER
			MAIL DATE 03/26/2009	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

RECEIVED

MAR 30 2009

ROBINS & PASTERNAK LLP

8325-0002.04

DOCKETED D8P

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte GEORGE NORBERT COX III,
CASEY CHRISTOPHER CASE, STEPHEN P. EISENBERG,
ERIC EDWARD JARVIS, and SHARON KAYE SPRATT

Appeal 2008-4125
Application 10/984,304
Technology Center 1600

Decided:¹ March 25, 2009 ✓

Before TONI R. SCHEINER, DONALD E. ADAMS, and
FRANCISCO C. PRATS, *Administrative Patent Judges*.

PRATS, *Administrative Patent Judge*.

¹ The two-month time period for filing an appeal or commencing a civil action, as recited in 37 C.F.R. § 1.304, begins to run from the decided date shown on this page of the decision. The time period does not run from the Mail Date (paper delivery) or Notification Date (electronic delivery). ✓

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134 involving claims to a method for modulating expression of a viral gene in a virally infected cell. We have jurisdiction under 35 U.S.C. § 6(b).

In a decision entered September 26, 2008 ("Decision"), we affirmed the Examiner's rejection of claims 70-72 under 35 U.S.C. § 103 as being obvious in view of Saiga and Nasser, and also affirmed the Examiner's provisional rejection of claims 70-72 for obviousness-type double patenting over copending Application No. 11/148,794 (Decision 17-18). We also remanded the case to the Examiner to reconsider an appealed rejection for indefiniteness under 35 U.S.C. § 112, second paragraph (*id.* at 18).

Because we remanded the case, we held the finality of our affirmances of the obviousness and obviousness-type double patenting rejections in abeyance until the proceedings on remand before the Examiner were concluded. 37 C.F.R. § 41.50(e).

On January 8, 2009, the Examiner informed Appellants that "[a]fter further consideration, the rejection of claims 70-72 under 35 U.S.C. 112, second paragraph is withdrawn" (Miscellaneous Communication 1 (January 8, 2009)). The Examiner's communication also stated that the case had been forwarded to the Board of Patent Appeals and Interferences (*id.*).

The case is now before us, with the only pending rejections being (1) the rejection of claims 70-72 under 35 U.S.C. § 103 as being obvious in view of Saiga and Nasser, and (2) the provisional rejection of claims 70-72 for obviousness-type double patenting over copending Application No. 11/148,794. For the reasons set forth in the Decision of September 26, 2009,

Appeal 2008-4125
Application 10/984,304

we affirm those rejections (Decision 3-15), and hereby make our decision FINAL.

TIME PERIOD

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a).

AFFIRMED

mls

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO CA 94303

The opinion in support of the decision being entered today was not written for publication and is not binding precedent of the Board.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Ex parte STEPHEN P. EISENBERG, CASEY C. CASE
GEORGE N. COX III, ANDREW JAMIESON
and EDWARD J. REBAR

Appeal No. 2006-0189
Application No. 09/825,242

ON BRIEF



Before MILLS, GRIMES, and GREEN, Administrative Patent Judges.

GRIMES, Administrative Patent Judge.

DECISION ON APPEAL

This appeal involves claims to a method of designing and making a zinc finger protein that binds to a selected target site, which the examiner has rejected as obvious. We have jurisdiction under 35 U.S.C. § 134. We reverse.

Background

"Zinc finger proteins (ZFPs) are proteins that can bind to DNA in a sequence-specific manner. . . . A single finger domain is about 30 amino acids in length. . . . To date, over 10,000 zinc finger sequences have been identified in several thousand known or putative transcription factors." Specification, page 1.

"The x-ray crystal structure of Zif268, a three-finger domain from a murine transcription factor, has been solved in complex with a cognate DNA-sequence. . . . The structure suggests that each finger interacts independently with DNA over 3 base-pair intervals, with side-chains at positions -1, 2, 3 and 6 on each recognition helix making contacts with their respective DNA triplet subsites." Id.

"The structure of the Zif268-DNA complex also suggested that the DNA sequence specificity of a zinc finger protein might be altered by making amino acid substitutions at the four helix positions (-1, 2, 3 and 6) on each of the zinc finger recognition helices." Page 2.

The specification describes "methods of designing zinc finger proteins that bind to a preselected target site. . . . The methods of design use a database containing information about previously characterized zinc finger proteins. This information includes names or other designations of previously characterized zinc finger proteins, the amino acid sequence of their component fingers, and the nucleotide triplets bound by each finger of the proteins. Information in the database is accessed using an algorithm that allows one to select fingers from different previous designs for combination in a novel zinc finger protein having specificity for a chosen target site."

Page 13.

Discussion

1. Claim construction

Claims 35, 40, 48, 49, and 53 are on appeal. Claims 37, 38, 41-43, and 52 are also pending; the examiner has indicated that these claims are allowable (Examiner's Answer, page 2).

Claim 35 is representative and reads as follows:

35. A method of producing a zinc finger protein or a nucleic acid encoding the same, comprising:

(a) providing a database comprising designations for a plurality of precharacterized zinc finger proteins, each protein comprising at least first, second and third fingers, and subdesignations for each of the three fingers of each of the zinc finger proteins, wherein at least one protein in the database has a third finger that is different from the third finger of at least one other protein in the database; a corresponding nucleic acid sequence for each zinc finger protein, each sequence comprising at least first, second and third triplets specifically bound by the at least first, second and third fingers respectively in each zinc finger protein, the first, second and third triplets being arranged in the nucleic acid sequence (3'-5') in the same respective order as the first, second and third fingers are arranged in the zinc finger protein (N-terminal to C-terminal);

(b) providing a preselected target site for design of a zinc finger protein, the target site comprising contiguous first, second and third triplets in a 3'-5' order,

(c) for the first, second and third triplet in the target site, identifying first, second and third sets of zinc finger protein(s) in the database, the first set comprising zinc finger protein(s) comprising a finger specifically binding to the first triplet in the target site, the second set comprising zinc finger protein(s) comprising a finger specifically binding to the second triplet in the target site, the third set comprising zinc finger protein(s) comprising a finger specifically binding to the third triplet in the target site;

(d) outputting designations and subdesignations of the zinc finger proteins in the first, second, and third sets identified in step (c); and

(e) producing (i) a zinc finger protein that binds to the target site comprising a first finger from a zinc finger protein from the first set, a second finger from a zinc finger protein from the second set, and a third finger from a zinc finger protein from the third set, or (ii) a nucleic acid encoding the zinc finger protein.

Thus, claim 35 is directed to a method of making a zinc finger protein that has been designed to bind to a specific target site. The method comprises providing a database that includes designations for zinc finger proteins that each comprise at least three zinc fingers. The claim requires that the database also include subdesignations for each of the three zinc fingers in each protein, as well as the complete (9-nucleotide)

nucleic acid sequence bound by the three zinc fingers in each of the zinc finger proteins.

Claim 35 requires the further steps of providing a target site for the zinc finger protein being designed (the target site being of a size that will be bound by three zinc fingers); identifying three sets of zinc finger proteins in the database, each of which includes proteins that include a zinc finger that will bind to one-third of the intended target site; outputting the designations for the proteins and zinc fingers in each of the three identified sets; and producing a zinc finger protein that includes the zinc fingers identified as binding to part of the intended target site.

2. Obviousness

The examiner rejected claims 35, 40, 48, 49, and 53 under 35 U.S.C. § 103 as obvious in view of Choo (PNAS),¹ Choo (Nature)² and Corbi.³ The examiner characterized Choo (Nature) as

show[ing] a method of designing a zinc finger protein that binds to a BCR-ABL recombined oncogene target site. . . . [E]ach triplet of the intended binding site (shown in figure 1) was used to screen a randomized zinc finger library. . . . Selected zinc fingers that bound a desired triplet were combined into a set of three finger zinc finger proteins.

Examiner's Answer, page 4. The examiner acknowledged that Choo (Nature) "does not show the extent of precharacterization of the zinc finger proteins in the randomized zinc finger library or a database of the randomized zinc finger library." Id., page 5.

¹ Choo et al., "Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage," Proc. Natl. Acad. Sci. USA, Vol. 91, pp. 11163-11167 (1994).

² Choo et al., "In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence," Nature, Vol. 372, pp. 642-645 (1994).

³ Corbi et al., "Synthesis of a new zinc finger peptide; comparison of its 'code' deduced and 'CASTing' derived binding sites," FEBS Letters, Vol. 417, pp. 71-74 (1997).

The examiner relied on Choo (PNAS) for teaching "a phage library comprising zinc finger genes in which the middle of three fingers is randomized. . . [and] show[ing] in figure 2 a database of selected and characterized library members." Id. The examiner relied on Corbi for its teaching of a zinc finger protein having three fingers, and the target site bound by the protein. Id.

The examiner concluded that

[i]t would have been obvious to a person of ordinary skill in the art at the time the invention was made to precharacterize the selected random library members of Choo et al. [(Nature)] to any desired extent to aid in further analysis of selected library members because Choo et al. [(PNAS)] shows such analysis in figure 2 and pages 11164-11167. It would have been further obvious to record such characterizations in a database as shown in Choo et al. [(PNAS)] figure 2. . . . It would have been further obvious to add other known zinc finger proteins and their binding specificities such as the Mago zinc finger protein of Corbi et al. to further increase the diversity of choices available in the database.

Id., page 6.

Appellants argue that, "even assuming arguendo that the cited references are properly combined, the references neither individually or in combination provide any disclosure of a database comprising designations for a plurality of three-finger zinc finger proteins, subdesignations for each of three fingers for each zinc finger protein, and their corresponding target nucleic acid sequences, as specified in claim[s] 35, 50, 48, 49, and 53." Appeal Brief, page 6 - 7. Appellants reason as follows:

The tables shown in Figs. 2 of the respective Choo references . . . provide designations for only a single finger of a multi-finger zinc finger protein and neither presents a target sequence with three triplets. Although the physical zinc finger proteins, from which the information discussed in the cited references was obtained, may inherently have had three zinc fingers, these physical proteins are not components of a database. If Fig. 2 of either Choo reference is viewed as a database, then the database is composed of the typewritten data in the tables. These typewritten data do

not expressly or inherently contain designations of zinc finger proteins, subdesignations of each of three fingers for each zinc finger protein, or the target sequences of the zinc finger proteins, as claimed. Corbi does nothing to compensate for the[se] deficiencies.

Id., page 7

We agree with Appellants that the cited references do not support a prima facie case of obviousness.

"In rejecting claims under 35 U.S.C. § 103, the examiner bears the initial burden of presenting a prima facie case of obviousness. Only if that burden is met, does the burden of coming forward with evidence or argument shift to the applicant."

In re Rijckaert, 9 F.3d 1531, 1532, 28 USPQ2d 1955, 1956 (Fed. Cir. 1993).

The test of obviousness is "whether the teachings of the prior art, taken as a whole, would have made obvious the claimed invention." In re Gorman, 933 F.2d 982, 986, 18 USPQ2d 1885, 1888 (Fed. Cir. 1991). The claimed invention is the invention defined by all the limitations of the claims. See In re Ochiai, 71 F.3d 1565, 1572, 37 USPQ2d 1127, 1133 (Fed. Cir. 1995) (A proper § 103 analysis requires "a searching comparison of the claimed invention – including all its limitations – with the teaching of the prior art.").

Here, the claimed method includes, among other things, "providing a database" comprising certain information. Thus, to make out a prima facie case of obviousness, the examiner must show that the prior art would have suggested providing a database that includes the information recited in the claims.

As Appellants point out, the database represented by Figure 2 of Choo (PNAS) does not include all of the information recited in the claims on appeal. The table shows the amino acid sequences of zinc finger regions of fusion proteins based on phage fd coat protein pIII (see Figure 1). Figure 2 of Choo (PNAS) also shows the trinucleotide sequences that are bound by each of the zinc fingers. However, Choo (PNAS) does not provide designations for complete zinc finger proteins that include the zinc fingers shown in Figure 2 or the nucleic acid sequences bound by the complete protein, which must include at least three zinc finger regions to meet the limitations of the claims on appeal. The other cited references also lack a teaching or suggestion of a database including all of the limitations recited in the instant claims.

The examiner responded to Appellants' argument on this point as follows:

Choo et al. [(PNAS)] makes clear in the second column of page 11163 that the library contains variants of a three finger Zif268 protein, and the tables show relevant characteristics of the members of the library.

Although Choo et al. [(PNAS)] does not list the entire sequence of each zinc finger protein in the table, the table is nevertheless a table of precharacterized three finger zinc finger proteins as claimed.

Examiner's Answer, page 6.

As we understand it, the examiner's position is that Choo (PNAS) discloses that a phage display library was made that contained a domain made up of three zinc fingers, the middle one of which was varied in order to make fusion proteins that would bind to different trinucleotide sequences, and Table 2 of the reference shows the amino acid sequence of the variant (middle) zinc finger. Thus, the examiner argues, although the table only shows the single, variant zinc finger, the skilled artisan would understand

it to represent an entire zinc finger protein because the rest of the amino acid sequence of each fusion protein was the same.

We agree with the examiner that those skilled in the art would have understood Table 2 of Choo (PNAS) to represent fusion proteins comprising three zinc finger domains: the variant shown in the table and two others that were common to all the members of the library. However, we do not agree that that understanding overcomes the deficiency of the reference.

All of the claims on appeal require the use of a database, not a library of physical compounds, that includes designations for zinc finger proteins that include three zinc finger domains, and "subdesignations for each of the three fingers of each of the zinc finger proteins," as well as the nucleic acid sequence bound by each of the zinc finger proteins (i.e., the nine-nucleotide long sequence bound by the set of three zinc fingers).

The database disclosed by Choo (PNAS) includes only the information relevant to the middle, variant zinc finger. It does not include subdesignations for the two invariant zinc fingers in each fusion protein, nor does it include the full nucleotide sequence bound by each of the fusion proteins. The cited references therefore do not teach all of the limitations of the claimed method, and the examiner has not adequately explained why a person of ordinary skill in the art would have found it obvious to modify the table/database of Choo (PNAS) to include the additional information.

Summary

The examiner has not shown that the claimed method, including all its limitations, would have been obvious to a person of ordinary skill in the art, based on the cited references. We therefore reverse the rejection of claims 35, 40, 48, 49, and 53.

REVERSED

Demetra J. Mills
Demetra J. Mills
Administrative Patent Judge

Eric Grimes
Eric Grimes
Administrative Patent Judge

Lora M. Green
Lora M. Green
Administrative Patent Judge

) BOARD OF PATENT
) APPEALS AND
) INTERFERENCES

Townsend and Townsend and Crew, LLP
Two Embarcadero Center
Eighth Floor
San Francisco, CA 94111-3834

The opinion in support of the decision being entered
today is *not* binding precedent of the Board.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte YEN CHOO, AARON KLUG, and MARK ISALAN

Appeal 2007-0743
Application 09/424,482
Technology Center 1600

Decided: August 23, 2007

Before TONI R. SCHEINER, DONALD E. ADAMS, and
NANCY J. LINCK, *Administrative Patent Judges*.

LINCK, *Administrative Patent Judge*.

ADAMS, *Administrative Patent Judge*, dissenting.

DECISION ON APPEAL

This is a 35 U.S.C. § 134 appeal in the above-referenced case.¹

We have jurisdiction under 35 U.S.C. § 6(b). We affirm.

¹ The application was filed on February 29, 2000. The real party in interest is Gendaq, Ltd., a wholly owned subsidiary of Sangamo Biosciences, Inc.

STATEMENT OF THE CASE

"The present claims are directed to libraries of zinc finger proteins in which positions 2 and 6 of adjacent zinc fingers are at least partially randomized." (Appellants' Revised Brief Under 37 C.F.R. § 41.37 (received March 20, 2006) (hereafter "Br.") 3.) According to the Specification,

Many DNA-binding proteins contain independently folded domains for the recognition of DNA, and these domains in turn belong to a large number of structural families, such as the leucine zipper, the "helix-turn-helix" and zinc finger families.

Despite the great variety of structural domains, the specificity of the interactions observed to date between protein and DNA most often derives from the complementarity of the surfaces of a protein α -helix and the major groove of DNA [Klug, (1993) Gene 135:83-92]. In light of the recurring physical interaction of α -helix and major groove, the tantalising possibility arises that the contacts between particular amino acids and DNA bases could be described by a simple set of rules; in effect a stereochemical recognition code which relates protein primary structure to binding-site sequence preference.

(Specification (hereafter "Spec.") 1.²) According to the Specification, no prior art method has succeeded in providing a complete code. (*Id.* at 2.)

Appellants based their library on the understanding that "zinc finger binding sites are determined by overlapping 4 bp subsites, and that sequence-specificity at the boundary between subsites arises from synergy between adjacent fingers." (*Id.* at 2-3.) Knowledge of this overlap resulted in Appellants' claimed invention which "provides a zinc finger polypeptide

² The numbers referenced for the Specification are those typed at the top of the pages and not the hand-written numbers at the bottom.

library in which each polypeptide comprises more than one zinc finger which has been at least partially randomised." (*Id.* at 3.) Consistent with their knowledge of this overlap, Appellants originally claimed: "A zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger and wherein each polypeptide has been at least partially randomized such that the randomization extends to cover the overlap of a single pair of zinc fingers." (Spec. 39.)

The claimed subject matter is reflected in representative claims 1 and 30, claims which restrict randomization at position 2 or 6 respectively.

These claims now read:³

1. A zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger comprising amino acid positions -1 to + 9 with position 1 representing the first amino acid of an alpha-helix and wherein each polypeptide has been at least partially randomised such that the randomisation extends to cover at least positions 6 and 2 of adjacent first and second fingers, respectively, wherein the randomisation of amino acid residues at position 2 is restricted to amino acids selected from the group consisting of D, A, R, Q, H, K, S, and N.

30. A zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger comprising amino acid positions -1 to + 9 with position 1 representing the first amino acid of an alpha-helix and wherein each polypeptide has been at least partially randomised such that the randomisation extends to cover at least positions 6 and 2 of adjacent first and second fingers, respectively, wherein the randomisation of amino acid at position 6 is restricted to amino acids selected from the group consisting of R, Q, V, A, E, K, N, and T.

³ Claims 1 and 30 are the only independent claims before us.

The Examiner has rejected claims 1-2, 6-7, and 27-34 under 35 U.S.C. § 103(a) over the following references:

Greisman, H.A. "*A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites*," 275 Science 657-61 (January 31, 1997) (hereafter "Greisman").

Choo, Y., "*Designing DNA-binding proteins on the surface of filamentous phage*", 6 Current Opinion in Biotechnology 431-36 (1995) (hereafter "Choo '95").

The Examiner also has rejected claims 1-2, 6, and 27-34 under 35 U.S.C. § 112, ¶ 1, for lack of written description.⁴

In addition to the above-cited references, the following additional references show the state of the art and thus the level of skill in the art:⁵

Choo & Klug, *Toward a code for the interactions of zinc fingers with DNA: selection of randomized zinc fingers displayed on phage*, 91 Proc. Natl. Sci. USA 11163-11167 (1994) (hereafter "Choo '94-1") (cited in Choo '95 and relied upon by Appellants (Br. 7 n.1 & accompanying text)).

Choo & Klug, *Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions*, 91 Proc. Natl. Sci. USA 11168-11172 (1994) (hereafter "Choo '94-2") (cited in Choo '95).

Choo & Klug, *Physical basis of a protein-DNA recognition code*, 7 Current Opinion in Structural Biology 117-25 (Feb. 1997) (hereafter "Choo '97").

Isalan, Choo and Klug, *Synergy between adjacent zinc fingers in sequence-specific DNA recognition*, 94 Proc. Natl. Acad. Sci. USA 5617-21 (May 1997) (hereafter "Isalan").

⁴ On appeal, the Examiner withdrew a § 112, ¶ 2 rejection of claims 6 and 32 (Answer 2) and admitted claim 7 should not have been rejected under § 112, ¶ 1 (Answer 3).

⁵ All of these references describe one or more of the inventors' own prior art work and were cited to the Office by the inventors. (See Information Disclosure Statements received Nov. 23, 1999 and Dec. 31, 2001.) In addition, two were cited in Choo '95.

OBVIOUSNESS UNDER § 103(a)

The § 103(a) Issue

Appellants contend:

[T]he cited references do not teach or suggest the groups of amino acids recited in independent claim 1 and claim 30 to which partial randomization of amino acids is restricted (D, A, R, Q, H, K, S, and N, in claim 1 and R, Q, V, A, E, K, N, and T in claim 30). . . . Greisman teaches to randomize in such a manner as to allow any of sixteen amino acids at each position (footnote 15 of Greisman). . . . Choo ['95] does not explicitly disclose whether he restricted randomization to certain codons. However, reference to his earlier work [Choo '94-1] . . . shows that he, like Greisman, restricted randomization to sixteen amino acids.

(Br. 7-8.)

The Examiner responds:

Greisman discloses at least at page 659, Fig. 3[A], the different zinc finger polypeptide[s] with random amino acid residues at position -1 up to position 6

Appellants recognized that Greisman teaches random amino acids with any of sixteen amino acids at each position omitting Trp, Phe, Tyr and Cys. [Note the instant claim does not also recite these residues, unless of course this is the non-preferred residues]. It would be within the ordinary skill in the art at the time the invention was made to pick and choose from the known available sixteen amino acids disclosed by Greisman, the ones that can combine to form the instant library. Greisman discloses, or at least suggests, from the 16 amino acids the random species at each position[] of the zinc finger as shown at Fig. 3. . . . Choo ['95] discloses that combinations or en bloc residues of zinc fingers can be random. Therefore, the combined teachings of the prior art would have led one having ordinary skill in the art at the time the invention was made to the instant random library. . . .

Appellants state that restricting the sets of amino acids at positions 2 and 6 is optimal . . .

In reply, neither claim 1 nor claim 30 restricts the sets of amino acids at positions 2 and 6. . . . Furthermore, the specification does not disclose that the restricted set of amino acids result[s] in an optimal library. Rather, [it discloses] that the library is screened for the residues that bind[] to the DNA target site. Be [that] as it may, optimization of a given parameter is within the ordinary skill in the art.

(Examiner's Answer (mailed June 1, 2006) (hereafter "Answer") 9-12 (first bracketed material in original).)

In response, Appellants point out:

[T]he zinc finger proteins whose sequences are shown in Fig. 3A of Greisman are not themselves a library of zinc finger proteins but rather eight individual clones from a library of zinc finger proteins (see, e.g., Greisman, p. 658, paragraph bridging cols. 1 and 2). . . . The actual physical zinc finger proteins existed as individual isolates and together with other zinc finger proteins as a large library, but not together as a library consisting only of the zinc finger proteins whose sequences are shown in Fig. 3A. . . .

Even if . . . it were assumed that the sequences of zinc finger proteins shown in Fig. 3A of Greisman were a library of zinc finger proteins, the diversity required by claim[] 1 . . . would still not be present. Claim 1 specifies inter alia that . . . randomization at position 2 is restricted to amino acids selected from the group D, A, R, Q, H, K, S and N. However, the randomization at position 2 is not so restricted in the zinc finger proteins shown in Fig. 3A. For example, position 2 of finger 1 contains a T in some of the sequences shown in Fig. 3A, a residue not allowed by the recited Markush group in claim 1.

(Reply Br. 4-5.)

With respect to the Examiner's position that it would have been obvious to optimize the amino acids, Appellants respond:

[S]election of subsets of amino acids for randomization to produce a library of zinc finger protein is not a continuous linear parameter that allows simple interpolations [like the situation in *In re Aller*]. Each of the amino acids is different, and the effect of its presence or absence at a given zinc finger position on the binding characteristics of a library of zinc finger proteins was unpredictable. Thus, selection of a set of amino acids to improve binding characteristics of a zinc finger protein library is not comparable to selecting a temperature or concentration of acid to conduct a chemical process.

Moreover, a *prima facie* case of obviousness based on optimization of ranges can be rebutted by showing that the art in any material respect teaches away from the claimed invention (*In re Geisler*, 43 USPQ2d 1362, 1366 (Fed. Cir. 1997)). Here, as discussed above, the art teaches that more different amino acids than were used by Greisman, not fewer (as claimed), would be desirable. The suggestion that more rather than fewer amino acids are desirable would have taught away from selecting a subset of the sixteen amino acids of Greisman.

(Reply Br. 8.)

We frame the § 103(a) issue: Would the zinc finger library of claims 1 and 30, including randomization in positions 6 and 2 of adjacent first and second fingers and restricted randomization in position 2 (claim 1) or 6 (claim 30), have been obvious to a skilled artisan, in view of Greisman's and Choo '95's teachings?

Findings of Fact Relating to Obviousness

Claim Interpretation

1. Giving the claim its broadest reasonable interpretation in light of the Specification, "library" requires two polypeptides. (Spec. 3 ("library is used . . . to denote a collection of polypeptides").)
2. "Randomization," as used in the claims and defined in the specification, "refers to the variation of the sequence of the polypeptides which comprise the library, such that various amino acids may be present at any given position in different polypeptides" (Spec. 6); but is not limited to simultaneous randomization on adjacent zinc fingers.
3. With respect to "each polypeptide," claim 1 requires randomization at "positions 6 and 2 of adjacent first and second fingers, respectively," and restriction of randomization at position 2 to "D, A, R, Q, H, K, S, and N."
4. Claim 30 requires randomization at "positions 6 and 2 of adjacent first and second fingers, respectively," and restriction of randomization at position 6 to "R, Q, V, A, E, K, N, and T."
5. Thus, claims 1 and 30 require a library of 2 polypeptides comprising 2 adjacent zinc fingers that have been partially randomized at positions 6 and 2, with randomization restricted to 8 of the 20 possible amino acids at a single position, i.e., position 2 (claim 1) or at position 6 (claim 30).

The Cited Prior Art

6. Greisman recognizes "the overlapping base contacts that can occur at the junction of neighboring subsites," i.e., positions 2 and 6 (Fig. 1B & Fig. 2 caption) and designs his strategy accordingly. (*Id.*)

7. Greisman identifies zinc finger proteins with high affinity for DNA target sites by recognizing “context-dependent interactions are important for zinc finger-DNA recognition” and “ensur[ing] that the new fingers are always selected in a relevant structural context.” (Greisman 657, cols. 2, 3.)

8. Greisman’s strategy involves a “method . . . for selecting DNA-binding proteins,” i.e., zinc finger proteins, by “adding and optimizing one finger at a time.” (Greisman 657 (abstract)).

9. Greisman uses the Zif268 structure, a known zinc finger structure, as a starting point and then “randomize[s] six potential base-contacting positions in each finger,” i.e., positions -1, 1, 2, 3, 5, and 6 (Greisman 657, col. 2 & 658 (Fig. 1)) using 16 amino acids and omitting 4 amino acids also omitted by Appellants. (Greisman note 15 & accompanying text.)

10. Greisman’s Fig. 2 illustrates their “protocol that successively selects finger 1, finger 2, and finger 3 to create a new zinc finger protein.” (Greisman 658 (Fig. 2 caption).)

11. Greisman begins with two known zinc fingers, Zif1 and Zif2 (wild-type Zif268 fingers) “to position the library of randomized fingers over the target site,” and then adds third zinc fingers from the randomized finger library, creating a “finger 1 library,” so called because this step selects potential finger 1 candidates for the ultimate 3-finger proteins. (Greisman 657, col. 3 & 658 (Fig. 2A).) From these potential “finger 1” candidates, Greisman selects and amplifies 3-finger peptides that bind to identified DNA binding sites, e.g., the TATA box. (*See id.*)

12. In step two, Greisman removes Zif1 from the 3-finger peptides but retains two of the fingers, Zif2 and select “finger 1,” and then ligates a “randomized finger 2 cassette” to these 2-finger peptides, again creating a

library from which 3-finger peptides that bind to identified DNA binding sites are selected and amplified. (*Id.* at 657, col. 3 & 658 (Fig. 2B).)

13. Finally, in step three, Greisman removes Zif2, leaving randomized fingers 1 and 2 that have been selected for their binding affinities, adds a “randomized finger 3 cassette,” and optimizes the final 3-zinc finger proteins.

14. Greisman’s protocol “actually was designed so that a *sublibrary of successful zinc finger sequences* could be carried over from one selection step (Fig. 2, A or B) to the next” and the final third step “then selects for combinations of fingers that work well together.” (Greisman 660 n. 16 (emphasis added); see also Greisman 660 n. 19 (“Each set of proteins exhibits a clear gradient of sequencing diversity across the three fingers (Fig. 3”).))

15. Greisman’s sublibrary of successful zinc finger sequences from step 3 that bind to the TATA box are shown in Fig. 3A. (Greisman 659; FFs 10-14.)

16. One result of Greisman’s strategy is a group of “new zinc finger proteins that recognize . . . the TATA box.” (Fig. 3A & caption.) Fig. 3A discloses the amino acid sequences for eight new 3-finger proteins that were selected from the third step library based on their ability to bind to the TATA box. (*Id.*; FFs 10-15.)

17. In each of the eight new zinc finger proteins that bind to the TATA box, Greisman discloses randomized amino acid sequences for Fingers 1, 2, and 3 at positions 2 and 6. (Greisman 659 (Fig. 3A).)

18. Further, in each of Greisman’s eight new zinc finger proteins, all the randomized amino acids in position 6 are those permitted by claim 30;

and in position 2, eleven of the amino acids (out of twenty-four) are those permitted by claim 1. (Greisman 659 (Fig. 3A; *see also* Answer 10).)

19. Greisman's sublibrary includes at least two polypeptides satisfying the restricted randomization for both positions 2 and 6. (Fig. 3A (6th & 7th clones, fingers 2 and 3).) In fact, Greisman "overexpressed" the 6th clone and used it for binding studies. (*See* Fig. 3 legend.)

20. Based on their binding potential, one of ordinary skill in the art would recognize the value of using Greisman's select zinc finger proteins (identified in Fig. 3A) in a zinc finger library for a number of purposes, including as a starting point for further selection or "to allow selection of proteins with four, five, or six fingers or to allow optimization of zinc fingers fused to other DNA-binding domains." (Greisman 659, col. 3.)

21. Since Greisman's sublibrary includes polypeptides that satisfy claim 1, to the extent Appellants' claimed amino acid selection has been optimized for position 2, one of ordinary skill in the art would have been motivated to optimize Greisman's selected zinc fingers through further randomization, coupled with binding experiments, such as those conducted by Greisman, with a reasonable likelihood of identifying amino acids with high affinity for DNA.

22. Choo '95 is also interested in identifying zinc fingers that bind to DNA and describes the physical basis of a possible protein-DNA recognition code, identifying the positions on the zinc fingers that make contact with the DNA, i.e., positions -1, 3, and 6 on one strand and position 2 on the second strand. (Choo '95, 432.)

23. Following a description of prior art publications, Choo '95 states: "From the large database of results, elements of a code can be deduced that

describe DNA recognition by zinc fingers.” (Choo ’95 at 432 (citing to Choo ’94-2 for discussion of this code).)

24. Greisman and Choo ’95 randomized to 16 amino acids, omitting Cys, Phe, Tyr, and Trp to avoid stop codons (*see Reply Br. 8*); of these 16, Appellants randomized to 8 with respect to position 2 or 6, also omitting Cys, Phe, Tyr, and Trp. (See claims 1 and 30.)

Additional Findings of Fact

25. Zinc finger binding motifs are structures “well known to those in the art and defined in” a number of references. (Spec. 14.)

26. “Detailed methodology for phage display is known in the art and set forth in” a number of references, and “[v]ector systems and kits for phage display are available commercially.” (Spec. 6.)

27. A number of alternative methods for generating zinc finger libraries with randomized amino acids at positions -1 through 6 were known to those skilled in the art at the time the invention was made. (*See, e.g.*, Greisman at 657-58 & nn. 15-16; Choo ’95 at 432-34.)

28. Randomized zinc finger libraries were known in the art, and DNA recognition sites on the fingers (positions -1 to 6) were also known. (*See, e.g.*, Greisman *passim*; Choo ’95, 431-32.)

29. Inter-finger “synergism,” or “context-dependent interactions,” due to overlapping 4 bp subsites was well recognized by those skilled in the art. (*See, e.g.*, Greisman 657-58; Isalan 5620; Choo ’97, 117.)

30. The cocrystal structures of a number of DNA-zinc finger complexes, including ones with zinc fingers Zif268 and Trk, are known, and the amino acids which bind to DNA have been identified as those at

positions -1, 3, and 6 on one finger and position 2 on the adjacent finger. (*See, e.g.*, Choo '94-1, 11166, col. 2; Choo '95, 432; Choo '97, 120 (Figs. 3 & 4); Isalan 5617; Choo '98, 2.)

31. With respect to these binding positions, it was also recognized that only certain amino acids are found at those positions. (*See, e.g.*, Choo '94-1, 11166, col. 2.)

32. The skilled artisan would have been motivated to optimize the amino acids on zinc finger proteins for a given DNA sequence and would have had a reasonable expectation of doing so, given the limited number of possible amino acids and the extensive guidance in the art. (FFs 25-31.)

Discussion of the § 103(a) Issue

Based on our findings and those of the Examiner, we conclude the invention claimed in claims 1 and 30 would have been obvious to one of ordinary skill in the art at the time the invention was made.

The skilled artisan's knowledge of zinc finger protein-DNA complexes is extensive. (FFs 6-31.) At the time the claimed invention was made, the basis for Appellants' invention, i.e., "synergism" or "context-dependent interactions" between position 6 on one finger and position 2 on an adjacent finger also was well recognized in the art. (FFs 6, 7, 22, 29 & 30.) Further, libraries of zinc finger proteins with at least two adjacent zinc fingers in which at least positions 2 and 6 were partially randomized to 16 amino acids (avoiding termination codons by not including Cys, Phe, Tyr, and Trp) were known. (FFs 6, 22, & 24; *see also* FFs 29 & 30.) Thus, such libraries were in the hands of the public to further refine and build upon.

Appellants now claim libraries of zinc finger proteins in which at least position 2 or 6 is partially randomized to 8 amino acids instead of 16, while

the remaining positions can be randomized to an unspecified number of amino acids. (Claims 1 and 30.) Thus, each claim would provide exclusivity for libraries that can be identical to those in the public domain, except for the restricted randomization in one position (limited to 8 amino acids rather than 16). In essence, Appellants are claiming a relatively large subgenus of the 16 amino acid genus in the prior art.

The situation here is similar to that in *In re Petering*, 301 F.2d 676, 681, 133 USPQ 275, 280 (CCPA 1962) in which the prior art genus disclosed 20 compounds and a limited number of generic R groups in the formula, and applicant attempted to claim a species within that genus. The CCPA stated:

[I]t is not the mere number of compounds in this limited class, . . . but, rather, the total circumstances involved, including such factors as the limited number of variations for R, only two alternatives for Y or Z, no alternatives for other ring positions, and a large unchanging parent structural nucleus. With these circumstances in mind, it is our opinion that Karrer has described to those with ordinary skill in this art each of the various permutations here involved as fully as if he had drawn each structural formula or had written each name.

Id. In this case, the prior art *fully describes* all 16 amino acids of the genus. Thus, as in *Petering*, one of ordinary skill in the art would “envisage *each member*” of the genus, including the 8 found in Appellants’ subgenus. *Id.* (emphasis in original).

The fact that Appellants’ 8 amino acids *may* bind better to DNA is not sufficient to render their claims patentable. The choice of amino acids in zinc fingers is a result-effective variable known to influence their binding to DNA. As such, one skilled in the art would have been motivated to optimize

that choice to improve binding using routine experimentation, i.e., techniques well known in the art. (FF 21.) Cf., e.g., *In re Boesch*, 617 F.2d 272, 276, 205 USPQ 215, 219 (CCPA 1980) (“discovery of an optimum value of a result effective variable in a known process is ordinarily within the skill of the art”); *Merck & Co. v. Biocraft Labs. Inc.*, 874 F.2d 804, 809, 10 USPQ2d 1843, 1847-48 (Fed. Cir. 1989) (“predictable results” obtained through “routine procedures” not sufficient to support validity). (See FFs 11-14, 24, & 26-28; Answer 12.) Further, to the extent selection was “unpredictable,” as Appellants argue (Reply Br. 8), unpredictability cannot be equated to nonobviousness when only a finite number of choices are available, as is the case here. See *Pfizer, Inc. v. Apotex, Inc.*, 480 F.3d 1348, 1364, 82 USPQ2d 1321, 1332 (Fed. Cir. 2007) (“obviousness cannot be avoided simply by a showing of some degree of unpredictability in the art so long as there was a reasonable probability of success”).

Appellants do not disclose any unexpected results obtained by limiting position 2 or 6 to the amino acids in their subgenus. This is not surprising, since claims 1 and 30 are not limited to any specific DNA binding sequences. Optimization depends upon both the nature of the zinc finger amino acids and the DNA binding sequence. Thus, we conclude it would have been obvious to try to optimize the choice of amino acids from 16 to 8 using techniques known in the art, and success in doing so would have been anticipated.

Under *KSR*, it’s now apparent “obvious to try” may be an appropriate test in more situations than we previously contemplated. When there is motivation

to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp. If this leads to anticipated success, it is likely the product not of innovation but of ordinary skill and common sense. In that instance the fact that a combination was obvious to try might show that it was obvious under § 103.

KSR Int'l Co. v. Teleflex Inc., 127 S. Ct. 1727, 1742, 82 USPQ2d 1385, 1397 (2007). This reasoning is applicable here. The “problem” facing those in the art was to determine what zinc finger amino acids would have high affinity for DNA out of the 16 amino acids randomized in the prior art, and there were a number of known methodologies available to do so. The skilled artisan would have had reason to use these methodologies to narrow the number of amino acids from 16 to 8, such as was done by Greisman, with the reasonable expectation of success. Thus, Appellants’ elimination of 50% of the amino acids in a single position was “the product not of innovation but of ordinary skill and common sense,” *id.*, leading us to conclude Appellants’ claimed library of zinc finger polypeptides claimed in claims 1 and 30 is not patentable.

Further, with respect to claim 30, a library of at least two polypeptides in which position 6 is randomized to R, Q, V, A, E, K, N, or T is fully disclosed by Greisman. (*See* FFs 11-19.) While Appellants argue Greisman’s Fig. 3A contains amino acids not permitted in position 2, they make no such argument with respect to position 6. Thus, we conclude claim 30 would have been obvious to the skilled artisan for this additional reason. *See, e.g., In re McDaniel*, 293 F.3d 1379, 1385, 63 USPQ2d 1462, 1466 (Fed. Cir. 2002) (“It is well settled that ‘anticipation is the epitome of obviousness.’”).

Appellants argue Greisman's zinc finger proteins disclosed in Fig. 3A are not a library. (Reply Br. 4.) In fact, Greisman teaches otherwise. (FFs 11-14.) Greisman creates a "sublibrary of successful zinc finger sequences" that bind to the TATA box. (FFs 14 & 16). This sublibrary meets the limitations of claim 30 and would have suggested that of claim 1 to the skilled artisan. (FFs 17-21; *see also* the discussion *supra* pp. 13-16.)

Appellants further argue Greisman "taught away from selecting a subset of the sixteen amino acids." (Reply Br. 8.) On the contrary, Greisman does exactly that, optimizing the zinc finger proteins disclosed in, e.g., Fig. 3A and thus limiting the number of amino acids in each position.

Dependent Claims 27, 28, 29, 33, and 34

Appellants separately argue claim 29 and claims 27, 28, 33, and 34.

Claim 29 is dependent upon claim 1 and restricts randomization at both positions 2 and 6 to those amino acids recited in claims 1 and 30.

Claims 27, 28, 33, and 34 are argued as a group. Thus, we select claim 27 as representative. Claim 27 restricts randomization at additional positions. It reads: "A library according to claim 1 wherein positions -1, 1, 2, 3, 5 and 6 of a first zinc finger and -1, 1, 2 and 3 of a second finger are randomized." These positions are those known to be "[k]ey base contacts." (*See, e.g.*, Greisman 658 (Fig. 1B & caption).)

We have considered Appellants' arguments relating to claims 27 and 29 and conclude their subject matter would have been obvious to one skilled in the art for the same reasons we so concluded with respect to claims 1 and 30. (*See "Discussion of the § 103(a) Issue," supra* pp. 13-16 & FFs 6-31.) Optimization of these binding sites would have been well within the level of skill in the art with routine experimentation.

WRITTEN DESCRIPTION UNDER § 112, ¶ 1

With respect to the written description issue, the Examiner finds the "specification, as originally filed, recites random *specific* amino acids for each of positions -1 to 6, particularly for the positions 6 and 2 pair," but does not support the language of claims 1 and 30, reciting random specific amino acids at position 2 or 6, respectively, but not at any other position. (Answer 3-4 (emphasis in original).)

In response, Appellants contend

the specification states at page 11, lines 13-14: "It is not necessary for each finger to be randomized at each of the positions . . . given in Table 1. In addition, the specification provides a table of amino acids "*preferably selected*" to appear at each position (p. 11, lines 3-14, emphasis supplied). The table lists D, A, R, Q, H, K, S, and N for position 2 and R, Q, V, A, E, K, N, and T for position 6. By using the term "*preferably selected*," the specification conveys that position 2 is preferably occupied by D, A, R, Q, H, K, S, and N, but can less preferably be occupied by other amino acids. Likewise position 6 is preferably occupied by R, Q, V, A, E, K, N, T, but can also be occupied by other amino acids. . . . The table does not state that if one position is occupied by a preferred group of amino acids, then every other position must also be occupied by its preferred group of amino acids.

(Br. 5.)

In view of the above, we frame the written description issue: Does Appellants' Specification, as filed, contain a written description sufficient to show they had possession of the full scope of their claimed invention at the time the application was filed, as required by Federal Circuit precedent?

Findings of Fact Relating to § 112, ¶ 1

33. As originally filed, claim 1 read: "A zinc finger polypeptide library in which each polypeptide comprises more than one zinc finger and wherein each polypeptide has been at least partially randomised such that the randomisation extends to cover the overlap of a single pair of zinc fingers." (Spec. 39.)

34. As originally filed, claim 7 read:

A library according to any preceding claim, wherein the randomisation of amino acid residues is restricted such that the following amino acids may appear at the given positions:

Position	Possible Amino Acids
-1	R, Q, H, N, D, A, T
1	S, R, K, N
2	D, A, R, Q, H, K, S, N
3	H, N, S, T, V, A, D
5	I, T, K
6	R, Q, V, A, E, K, N; T

(Spec. 39.)

35. According to the Specification: "It is not necessary for each finger to be randomised at each of the positions." (Spec. 11.)

36. Claims 1 and 30 are of intermediate scope between originally-filed claims 1 and 7 (see Answer 6), and thus are supported by the originally-filed claims which are part of the specification.

Discussion of the Written Description Issue

With respect to claims 1 and 30, we find the written description requirement of § 112, ¶ 1, is satisfied. (FFs 33-36.) The Examiner does not provide any separate bases to support her written description rejection of

dependent claims 2, 6, 27-29, and 31-34. Thus, we find the written description requirement also is satisfied for claims 1-2, 6, and 27-34.

CONCLUSION

In summary, we affirm the § 103(a) rejection of claims 1 and 30, and reverse the § 112, ¶ 1 rejection of claims 1-2, 6, and 27-34 based on lack of written description.

Pursuant to § 41.37(c)(1)(vii)(2006), we also affirm the rejection of claims 2, 6-7, 28, and 31-34 under § 103(a), as these claims were not argued separately.

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a)(1)(iv)(2006).

AFFIRMED

ADAMS, *Administrative Patent Judge*, dissenting.

I disagree with the majority's decision to affirm the rejection of claims 1, 2, 6, 7, 27-34 under 35 U.S.C. § 103. In my opinion, the evidence on this record lacks the necessary factual foundation to support this conclusion. Accordingly, I dissent.

Because the majority correctly finds that Appellants' Specification provides written descriptive support for claims 1-2, 6, and 27-34, I limit my discussion to the rejection under 35 U.S.C. § 103.

Overview:

The claims before us on appeal are drawn to a zinc finger polypeptide library wherein each polypeptide comprises more than one zinc finger (e.g., three zinc fingers⁶). "Zinc fingers, as is known in the art, are nucleic acid binding molecules" (Specification 6⁷: 23). The DNA binding portion of a zinc finger consists of an α -helix. The art has applied a numbering scheme for this α -helix, wherein the amino acid positions in the α -helix are numbered consecutively starting with the first amino acid of the α -helix as

⁶ While more than one zinc finger reads on at least two zinc fingers, the issue before this panel does not turn on the number of zinc fingers in each polypeptide. Therefore, I discuss the claimed zinc finger polypeptide library in the context of a library, wherein each polypeptide comprises three zinc fingers. I do so, because the prior art before this panel teaches zinc finger polypeptide libraries wherein each polypeptide comprises three zinc fingers (e.g., three finger libraries).

⁷ For clarity, I note that the pages of Appellants' Specification have page numbers at both the top and bottom of the page. Unfortunately, the page numbers on any given page do not correspond to each other. Accordingly, I will refer to those page numbers that appear at the top of each page of Appellants' Specification.

+1 (*see, e.g.*, Choo '94-1 11164: col. 2, ll. 5-7). Accordingly, the -1 position represents the amino acid immediately preceding the +1 position.

It has been found that due to the consecutive turns in the α -helix of the zinc finger; positions -1, 3, and 6 of each finger make a specific contact with a single DNA strand (*e.g.*, a first strand) at the 3' middle and 5' positions of a 3-base pair (bp) DNA target site (*see, e.g.*, Isalan 5617: col. 1, ll. 5-10; 5618: Fig. 1a). It has also been found that position 2 of the zinc finger may play a part in the recognition of DNA by binding the complimentary (*e.g.*, second) DNA strand. Specifically, while positions -1, 3, and 6 make contact with the DNA target site and are considered to be the primary amino acid positions in the zinc finger α -helix; it has been found that with polypeptides comprising at least two zinc fingers, an auxiliary position, the 2 position of the preceding zinc finger may also be involved in the recognition of DNA (Choo '94-1 11166: col. 2, ll. 67-69; Choo '97 117: col. 2, ll. 36-42).

However, the contribution of this "auxiliary" 2 position of a zinc finger to DNA recognition and binding is less than clear. For example, Choo '94-2 report that "[t]he amino acid at position +2 is able to modulate or enhance the specificity of the amino acid at other positions" (Choo '94-2 11170: col. 2, ll. 9-11). This same group, however, reports that but for a specific instance when the -1 position is the amino acid Arg and the amino acid at the +2 position is Asp, any contacts this +2 position "might make with the second DNA strand do not contribute significantly to the stability of the protein-DNA complex" (Choo '94-1 11166: col. 1, ll. 15-19).

Nevertheless, to account for any possible effect the +2 position of a zinc finger may have on DNA recognition and binding, when a zinc finger polypeptide comprising at least two zinc fingers is considered, the DNA

binding site expands from 3-bp to 4-bp (Isalian 5618: col. 1, Fig. d; 5620: col. 2, ll. 37-40). In this regard, “[e]ach zinc finger binds to a quadruplet sequence in a target nucleic acid through contacts between specific amino acid residues of the α -helix of the zinc finger and the nucleic acid strand” (Specification 6: 23-26). Appellants explain that “[t]he quadruplets specified in the present invention are overlapping, such that, when read 3’ to 5’ on the –strand of the nucleic acid, base 4 of the first quadruplet is base 1 of the second, and so on” (Specification 6: 26-28; Fig. 1).

Recognizing that, in certain circumstances, there is an interrelationship between two adjacent zinc fingers of a zinc finger polypeptide, those of ordinary skill in the art have appreciated that the evaluation of zinc fingers is best performed in what is referred to as “context-dependent interactions,” e.g., the interactions between neighboring fingers and DNA target sites (Greisman 657: col. 3, ll. 24-28). Therefore, the art before us on this record speaks of three finger libraries, e.g., libraries of zinc finger polypeptides that comprise three zinc fingers.

As outlined below, the work in this field has attempted to develop a set of rules through which one can predictably design a zinc finger polypeptide that will recognize and bind a particular DNA target site. The reports of this work in the literature, at the time Appellants’ claimed invention was made, indicate that while some interesting observations have been made; there is still “no general code that can be used to design optimal zinc finger proteins for any desired target sequence or that can predict the preferred binding site of every zinc finger protein” (Greisman 659: col. 2, ll. 16-21; *see also* Choo ’94-1 11167: col. 1, ll. 3-6 (“although sequence homologies are strongly suggestive of amino acid preferences for particular

base pairs, we cannot confidently deduce such rules until the specificity of individual fingers for DNA triplets is confirmed.”)

The work in this field has led those of ordinary skill in the art to the conclusion that the zinc finger recognition code is a degenerate code (Choo '97 121: col. 1, ll. 7-8). In this regard, those of ordinary skill in the art have found that the structural conformation of the zinc finger itself; and more particularly, the α -helices contribute to the degeneracy of the zinc finger code (Choo '97 121: col. 1, ll. 8-15). In addition, those of ordinary skill in this art have found that “[a]lthough the zinc finger structure directly enables the binding mode that results in the possibility of coded contacts, we must emphasize that protein-DNA recognition is mutual and that DNA is, therefore, not a passive participant in the binding reaction” (Choo '97 124: col. 1, ll. 26-30).

Further complicating the attempts in the art to develop a code, or set of rules, that can be used to predictably design zinc finger polypeptides is the recognition that even when one selects a naturally occurring zinc finger polypeptide (e.g., Zif268) as a starting point - “natural DNA-binding sites for proteins are rarely the optimal binding sequences, and that naturally occurring proteins often evolve to recognize a variety of different binding sites” (Choo '97 123: col. 1, ll. 1-5).

Nevertheless, despite the foregoing hurdles that must be overcome in order to predictably design zinc finger polypeptides, those in the field remain optimistic that the degenerate zinc finger code will be cracked, so - the work continued.

Appellants now come before this panel with a disclosure of a library “in which randomization is limited to substituting amino acids which are known to dictate variation in binding site specificity” (Specification 3:1-3). According to Appellants’

[t]he present invention provides a code of amino acid position bias which permits the selection of the library against any nucleic acid sequence as the target sequence, and the production of a specific nucleic acid-binding protein which will bind thereto. Moreover, the invention provides a method by which a zinc finger protein specific for any given nucleic acid sequence may be designed and optimized. The present invention therefore concerns a recognition bias which has been elucidated for the interactions of classical zinc fingers with nucleic acid. In this case a pattern of rules is provided which covers binding to all nucleic acid sequences.

(Specification 3: 23-30.)

Claim interpretation:

Claims 1 and 30 are the only independent claims of record. Claims 2, 6, 7, and 27-29 depend from claim 1; claims 31-34 depend from claim 30. The claims are drawn to a zinc finger polypeptide library. The claimed library is a composition, e.g., a product.

Notwithstanding the majority’s supposition⁸, as Appellants explain

⁸ In what they refer to as Finding of Fact #1, the majority asserts that a “library” requires two polypeptides. (Spec. 3 [sic, 5] (‘library is used . . . to denote a collection of polypeptides’)’ (*supra* 8). The majority’s definition of the term “library” lacks the precision to which those of ordinary skill in the art would recognize the term to represent. Under the majority’s definition a “library” would represent a composition consisting of two of the exact same polypeptides. This is not how a person of ordinary skill in the art would interpret this term.

[t]he term “library” is used according to its common usage in the art, to denote a collection of polypeptides or, preferably, nucleic acids encoding polypeptides. The polypeptides of the invention contain regions of randomization, such that each library will comprise or encode a repertoire of polypeptides, wherein individual polypeptides differ in sequence from each other.”

(Specification 5: 3-7.) Therefore, even if a library is construed to represent only two polypeptides, those two polypeptides must be in the same composition and must differ in sequence from each other to be considered a library.

The claimed zinc finger polypeptide library requires that each polypeptide of this library comprises more than one zinc finger (e.g., at least two zinc fingers; in the context of the prior art before us on this record - a library, wherein the zinc finger polypeptides have three zinc fingers). Each zinc finger comprises amino acid positions -1 to +9, wherein position 1 represents the first amino acid of an alpha helix. As discussed above, this numbering scheme is consistent with the numbering scheme used in the art.

If this was all Appellants’ had claimed, the zinc finger libraries would be anticipated by any number of prior art references, e.g., Greisman. Appellants, however, have not claimed a generic zinc finger library. To the contrary, Appellants have claimed a very specific zinc finger library. More specifically, Appellants’ library is defined by the process by which it is made.

The process that defines the claimed product:

According to Appellants' claims, the polypeptides of the claimed library are at least partially randomized, wherein the randomization⁹ extends to cover at least positions 6 and 2 of adjacent first and second fingers. Therefore, at a minimum, claims 1 and 30 require that those amino acids at positions 6 and 2 of the adjacent first and second fingers respectively are randomized. Appellants' Specification discloses this concept as a preferred embodiment of the invention (Specification 9: 6-7 "[p]referably, each library will comprise randomisation . . . [of] at least position 6 of the first finger and position 2 of a second finger."). For clarity, the following illustration characterizes a three finger library, wherein the zinc finger polypeptides have amino acid positions -1 to 9. The box around positions 6 and 2 emphasize the adjacent relationship of the only two amino acids of this zinc finger polypeptide that are absolutely required, by the claims on appeal, to be randomized.

1 st Finger	2 nd Finger	3 rd Finger
-1 1 2 3 4 5 6 7 8 9	-1 1 2 3 4 5 6 7 8 9	-1 1 2 3 4 5 6 7 8 9

The process does not end there. According to Appellants claims position 2 of the second finger and position 6 of the first finger are randomized with a specific set of eight amino acids. Specifically:

Claim 1 requires that the randomization of the amino acid residue at position 2 is restricted to amino acids selected from the group consisting of D, A, R, Q, H, K, S, and N.

⁹ According to Appellants the term "[r]andomisation, as used herein, refers to the variation of the sequence of the polypeptide which comprise the library, such that various amino acids may be present at any given position in different polypeptides" (Specification 5: 11-13).

Claim 30 requires that the randomization of the amino acid residue at position 6 is restricted to amino acids selected from the group consisting of R, Q, V, A, E, K, N, and T.

Summary:

The claimed zinc finger polypeptide library is a collection of zinc finger polypeptides (e.g. at least two zinc finger polypeptides) that differ in sequence from each other.

Each polypeptide in the library comprises at least two zinc fingers.

The amino acid sequences of the polypeptides in the library are at least partially randomized, but they must, at a minimum, be randomized at position 6 of the first finger and position 2 of the second finger. In addition, claims 1 and 30 set forth the following rules when randomizing the amino acids at positions 2 and 6:

According to claim 1, position 2 of the second finger must be an amino acid chosen from the group consisting of D, A, R, Q, H, K, S, or N. Position 6 of the first finger, however, can be randomized with any amino acid.

According to claim 30, position 6 of the first finger must be an amino acid chosen from the group consisting of R, Q, V, A, E, K, N, and T. In claim 30, position two can be randomized with any amino acid.

Issue:

The issue before this panel is whether the prior art, alone or in combination, teaches or suggests a zinc finger polypeptide library, wherein the amino acids residues at least at position 2 or 6 are randomized according

to the randomization rules set forth in Appellants' claims 1 and 30 respectively.

The Rejection of Record:

The claims stand rejected under 35 U.S.C. § 103 as being unpatentable over the combination of Greisman and Choo '95.

Findings of Fact (FF):

- i. Phe, Tyr, Trp, and Cys do not, or only rarely, appear in positions -1 to 8 of a zinc finger (Choo '94-1 11164: col. 2, ll. 45-49; Greisman 660 n. 15).
- ii. Greisman teach a zinc finger library wherein positions -1, 1, 2, 3, 5, and 6 were randomized with codons that "allow 16 side chains at each position (all residues except Cys, Phe, Tyr, and Trp) . . ." (Greisman 658: Fig. 1, legend; 660: col. 1, n. 15). Stated differently, Greisman produced a zinc finger library, wherein positions -1, 1, 2, 3, 5, and 6 of each finger in the library where randomized with 16 different amino acids.
- iii. Griesman report that "[a]ll phage display libraries contained between 5.6×10^8 and 1.9×10^9 clones" (Gresiman 660 n. 15).
- iv. While conserved residues often appear at position -1, 3, or 6 of the α helix when adenine or guanine is present in the primary strand of one of Greisman's binding sites; Greisman found no such simple patterns at other positions in their selected proteins. Specifically, Greisman report that they "found no simple patterns of residues at positions 1, 2, and 5 of the α helix, and when thymine or cytosine occurs on the primary strand (Fig. 3), we found no simple pattern of potential contacts from residues at positions

-1, 3, and 6" (Greisman 659: col. 1, l. 33 - col. 2, l. 5).

v. At the end of Greisman's study, the authors report that "there still is no general code that can be used to design optimal zinc finger proteins for any desired target sequence or that can predict the preferred binding site of every zinc finger protein" (Greisman 659: col. 2, ll. 16-21).

vi. While Choo '95 reports that "[s]eventeen different triplets were used in successful selection experiments" and that "[t]he zinc fingers selected by a given triplet were found to have a bias towards a particular amino acid in three positions (-1, 3 and 6) . . ." (Choo '95 432: col. 1, ll. 32-36); Choo '95, does *not* identify which amino acids the triplets were biased towards.

vii. "In almost all of the selected fingers in which Arg recognizes G[uanine] at the 3' end, Asp occurs at position +2 to buttress the long Arg side chain . . ." (Choo '94-1 11166: col. 1, ll. 13-15; Choo '97 119: col. 2, ll. 14-18 ("[i]n certain cases . . . two amino acids from different positions appear to cooperate in specifying one base . . .; for instance, the buttressing interaction of position 2, aspartate, with position -1, arginine, which contacts guanine . . .").

viii. "When position -1 is not Arg, Asp rarely occurs at +2, suggesting that in this case any other contacts it might make with the second DNA strand do not contribute significantly to the stability of the protein-DNA complex" (Choo '94-1 11166: col. 1, ll. 15-19).

ix. "When adenine is present at the 3' end of a triplet, Gln is often selected at position -1 of the α -helix, accompanied by small aliphatic residues at +2. . ." (Choo '94-1 11166: col. 2, ll. 3-6).

x. "Thymine at the 3' end of a triplet selects a variety of polar amino acids at -1 . . . and occasionally returns fingers with Ser at +2 . . ." (Choo '94-1 11166: col. 2, ll. 34-38).

xi. Based on "[p]reliminary modeling studies" Isalan speculates that "histidine in position 2 might make a cross-strand contact to G[uanine] or T[reonine] while maintaining the buttress to Arg -1." (Isalan 5620: col. 1, ll. 10-14). In this regard, Isalan reports that "phage selections from randomized C-terminal finger libraries have yielded several fingers with His2, and Leu or Ser at position 1 . . ." (Isalan 5620: col. 1, ll. 14-18).

xii. "Gln at position -1, which is specific for adenine at the 3' end of a triplet when position +2 is a small nonpolar amino acid such as Ala but is specific for thymine when a polar residue such as Ser is at position +2" (Choo '94-2 11170: col. 2, ll. 2-5).

xiii. "The amino acid at position +2 is able to modulate or enhance the specificity of the amino acid at other positions" (Choo '94-2 11170: col. 2, ll. 9-11).

xiv. In a three finger library, "Asp at position +2 of finger 3 is dominant over the amino acid present at position +6 of the middle finger, precluding the possibility of recognition of adenine or cytosine at the 5' position" (Choo '94-2 11170: col. 1, ll. 52-56).

xv. With regard to the Zif268 zinc finger polypeptide binding domain,

[t]he first and third fingers have positions: -1. arginine: 3. glutamate: and 6. arginine: and bind 5' -GCG-3'; whereas the middle finger has positions: -1. arginine: 3. histidine: and 6. threonine: and binds 5' -TGG- 3'.... Only one type of contact is observed to the second DNA strand, by position 2. aspartate,

which contacts a cytosine found in the subsite of the preceding zinc finger

(Choo '97 117: col. 1, ll. 36-42.)

xvi. When present at the 5' or 3' end of a triplet, guanine selects fingers with Arg at positions +6 or -1 of the α -helix, respectively (Choo '94-1 11166: col. 1, ll. 2-5 and 8-10).

xvii. "Occasionally, guanine at the 5' end of a triplet selects Ser or Thr at +6 . . ." (Choo '94-1 11166: col. 1, ll. 7-8).

xviii. "When thymine is at the 5' end of a triplet, Ser and Thr are selected at +6 (as is occasionally the case for guanine at the 5' end)" (Choo '94-1 11166: col. 2, ll. 32-34).

xix. "The triplets ACG . . . and ATG . . . , which have adenine at the 5' end, also returned oligoclonal mixtures of phage, the majority of which were of one clone with Asn at +6" (Choo '94-1 11166: col. 2, ll. 10-13).

xx. "Asp is also sometimes selected at +3 and +6 when cytosine is in the middle . . . and 5' . . . position respectively" (Choo '94-1 11166: col. 2, ll. 18-20).

xxi. "[I]n order that Ala may pick out thymine in the triplet GTG, Arg must not be used to recognize guanine from position +6, since this would distance the Ala residue too far from the DNA . . ." (Choo '94-2 11171: col. 2, ll. 30-35).

xxii. "Guanine bases in our sites appear to prefer Arg at positions -1 and 6 Adenine bases appear to . . . prefer Gln at position -1 and, to some extent, at position 6" (Greisman 660 n. 20).

xxiii. When

adenine or guanine occurs in the primary strand of one of our binding sites (the strand corresponding to the guanine-rich strand of the Zif268 site), there often is a conserved residue at position -1, 3, or 6 of the α helix that could form hydrogen bonds with this base . . .

(Greisman 659: col. 1, l. 20-26, endnote omitted.)

xxiv. “[I]t appears that the identity of an amino acid at any one α -helical position [e.g., the -1 position] is attuned to the identity of the residues at the other two positions [e.g., the 3 and 6 positions] to allow three base contacts to occur simultaneously” (Choo '94-2 11171: col. 2, ll. 27-30).

xxv. “[A]lthough sequence homologies are strongly suggestive of amino acid preferences for particular base pairs, we cannot confidently deduce such rules until the specificity of individual fingers for DNA triplets is confirmed” (Choo '94-1 11167: col. 1, ll. 3-6).

xxvi. Choo'97 points out that

[o]ne physical complication that leads to a degenerate code is that the helical pitch of the α -helix that means that the three recognition positions do not fall on exactly the same face of the helix. Hence, the three base-contacting positions are not equidistant from the DNA, and residues at position 3 of the helix (which is closer to the DNA) need to be shorter than at either positions -1 or 6.

Choo '97 121: col. 1, ll. 8-15; *see also* Choo '94-2 11171: col. 2, l. 35 – 11172: col. 1, l. 6.)

xxvii. Choo '97 provides a code table that illustrates that a number of different amino acids can specify a particular base in the DNA binding site (Choo '97 121: col. 1, ll. 1-2; Fig. 5). This exact same code table appears in

Choo '94-2 (Choo '94-2 11170: col. 2, Fig. 2). Choo '97 caution, however, that this table

should not imply that all combinations of amino acids will be effective. The amino acid used by zinc fingers to specify a particular base depends on the position of that base in the cognate triplet, and sometimes on its sequence in context of the triplet. The zinc finger recognition code is therefore a degenerate code.

(Choo '97 121: col. 1, ll. 3-8.)

Analysis:

The prior art before this panel teaches randomized three finger zinc finger polypeptide libraries. There is, however, no teaching or suggestion in any of this prior art, alone or in combination, that teaches a zinc finger polypeptide library that falls within the scope of the requirements set forth in Appellants' claimed invention.

Griesman produce a zinc finger library wherein positions -1, 1, 2, 3, 5, and 6 of each finger in the library are randomized with 16 different amino acids (FF ii). The only 4 amino acids that Griesman did not use where those four that the prior art recognized did not, or only rarely, appear in positions -1 through 8 of known zinc fingers (FF i).

Griesman report that “[a]ll phage display libraries contained between 5.6×10^8 and 1.9×10^9 clones” (FF iii). Griesman report that after “multiple rounds of selections (Fig. 2) were completed, the final phage pools [(libraries)] bound tightly to their respective target sites [(e.g., the TATA box, p53 binding site, the nuclear regulatory element)]. DNA sequencing of *eight clones* from each *pool* [(e.g., sublibrary)] revealed marked patterns of

conserved residues (Fig. 3) . . .” (Greisman 658: col. 2, l. 1 to col. 3, l. 3, emphasis added).¹⁰ Greisman report the sequencing data of the eight clones selected from each screen in Fig 3 (Greisman 659: Fig. 3). The data listed in Fig. 3 is simply a listing of the deduced amino acid sequences of clones isolated from a zinc finger library. There is no indication in Greisman that the sequence of any of these clones is representative of all the polypeptides in the library, or the sublibraries of zinc fingers having specificity for one of the DNA target sequences. More specifically, there is no indication in Greisman that the amino acids at position 2 or 6 of these clones are representative of the amino acids at position 2 or 6 in the remaining polypeptides of the library or sublibrary. Given that Greisman generated libraries by randomizing positions -1, 1, 2, 3, 5, and 6 with 16 different amino acids in each position (FF ii), there is no reasonable expectation that the library, or any of the sublibraries, would be expected to contain only those zinc finger polypeptides that fall within the scope of Appellant’s claims. Stated differently, there is no suggestion or teaching in Greisman of a library within the scope of Appellants’ claimed invention.

There is also no teaching or suggestion in Greisman that when a randomized zinc finger polypeptide library is produced, the choice of amino acid at any given position in the zinc finger polypeptide should be reduced or restricted. There is no teaching or suggestion in Greisman that position 2 of the second finger, or 6 of the first finger, of a zinc finger polypeptide

¹⁰ To dispel a myth that appears to have, at least inferentially, clouded both the Examiner’s and the majority’s reasoning; the clones that were selected for sequencing and reported in Fig. 3 of Greisman are *not* a library. They are individual clones – single polypeptides.

should be randomized only with the specific amino acids set forth in Appellants' claimed invention.

To the contrary, while Greisman find that under certain circumstances, a pattern may be found for positions -1, 3, and 6; there was no "simple patterns of residues at positions 1, 2, and 5 of the α helix . . ." (FF iv). In all, Greisman teaches a person of ordinary skill in the art that "there is no general code that can be used to design optimal zinc finger proteins for any desired target sequence or that can predict the preferred binding site of every zinc finger protein" (FF v). Therefore, it cannot be said that Greisman would have taught or suggested a zinc finger library within the scope of Appellants' claimed invention.

Choo '95 does not make up for the deficiencies in Greisman. Choo '95 states that they observed that "zinc fingers selected by a given triplet were found to have a bias towards a particular amino acid in three positions (-1, 3 and 6) . . ." (FF vi). Choo '95, however, provides no description of these triplets or the particular amino acid that these triplets reportedly preferred. Having failed to suggest a preference for any particular amino acid at an position in the zinc finger it cannot be said that Choo '95 would have taught or suggested a zinc finger library within the scope of Appellants' claimed invention.

On reflection, the facts on this record establish that neither Greisman nor Choo '95 teach or suggest a library that falls within the scope of Appellants' claimed invention. Accordingly, I would reverse the rejection of claims 1, 2, 6, 7, and 27-34 under 35 U.S.C. § 103 as being unpatentable over the combination of Greisman and Choo '95.

The Examiner's position:

The Examiner directs attention to Greisman's Fig. 3, in an attempt to demonstrate that Greisman's teachings would have lead a person of ordinary skill in the art to produce a zinc finger library wherein the amino acid residue at position 2 or 6 is randomized with the restricted population of amino acids set forth in Appellants' claimed invention (Answer 10). According to the Examiner, Greisman's Fig. 3A illustrates that the 2 and 6 positions of the 4th Finger 1 sequence are "N", the 2 position of the 6th-8th Finger 2 sequences are "A", and the 6 position of all the Finger 2 sequences are either Q or A (Answer 10). From this the Examiner concludes that the sequences set forth in Fig. 3A "meet the claimed library with at least position 2 being a random residue at one finger and the adjacent finger with 6 having a random residue" (*id.*). Based on this reasoning, the Examiner concludes that "the claimed library with the random amino acid residues includes or encompasses the library of random peptide of Greisman for each of the different fingers" (*id.*).

There are two problems with the Examiner's reasoning: First, as discussed above, the sequences illustrated in Greisman's Fig. 3 are of isolated clones – single polypeptides. These clones are not a library. Second, the Examiner's use of the phrase "includes or encompasses" suggests that the Examiner is of the opinion that as long as one of the amino acids listed in Appellants' claims for position 2 or 6 is found in a zinc finger polypeptide library, that library will read on the claim. This rationale is incorrect. According to Appellant's claims, the library is *restricted* to the amino acids "selected from the group consisting of . . ." (see claims 1 and 30). Stated differently, Appellants' claimed library is *restricted* to contain

only those zinc finger polypeptides that contain the amino acids listed in claims 1 and 30 at position 2 of the second finger and position 6 of the first finger respectively. Neither Greisman nor Choo '95 teach such a restricted library; nor do they suggest restricting the randomization of a zinc finger library in the manner set forth in Appellants' claims.

For the foregoing reason, I am not persuaded by the Examiner's rationale. The Examiner fails to adequately explain on this record why a person of ordinary skill in the art would have been lead to Appellants' claimed library based on the teachings of Greisman and Choo '95.

I recognize the Examiner's assertion that "[i]t would be within the ordinary skill in the art at the time the invention was made to pick and choose from the known available sixteen amino acids disclosed by Greisman, the ones that can combine to form the instant library." Stated differently, since Greisman teaches the randomization of positions -1, 1, 2, 3, 5, and 6 of each finger in the library with 16 different amino acids (FF ii); it would have been obvious to look at Appellants' claims and Specification, and select the same amino acids that Appellants selected. This assertion is the epitome of hindsight. As set forth in *KSR Int'l Co. v. Teleflex Inc.*, 127 S. Ct. 1727, 1742, 82 USPQ2d 1385, 1397 (2007),

A factfinder should be aware, of course, of the distortion caused by hindsight bias and must be cautious of arguments reliant upon *ex post* reasoning. See *Graham*, 383 U.S., at 36, 86 S.Ct. 684 (warning against a "temptation to read into the prior art the teachings of the invention in issue" and instructing courts to "guard against slipping into the use of hindsight" (quoting *Monroe Auto Equipment Co. v. Heckethorn Mfg. & Supply Co.*, 332 F.2d 406, 412 (C.A.6 1964))).

I recognize the Examiner's assertion that

Greisman discloses, or at least suggests, from the 16 amino acids the random species at each positions [sic] of the zinc finger as shown at Fig. 3. Some of the combination of amino acids includes a random residue at positions 2 and 6, which is encompassed by the instant random residues at said positions. Gresiman [sic] teaches that each of the amino acids in the zinc finger is random amino acid that can be selected from 16 amino acids and discloses some specific ones.

(Answer 10-11.) As best as I can understand it, the Examiner seems to be asserting that since Greisman teaches that some of the amino acid residues at positions 2 and 6 of the representative clones set forth in Fig. 3 are the same as those set forth in Appellants' claims a person of ordinary skill in the art would have found it to have been *prima facie* obvious at the time Appellants' invention was made to select only those that would have met the requirements of Appellants' claimed invention. The problem is, however, that there is no evidence on this record that leads to this conclusion. Particulary, when Greisman states that no preferences were found for any position in a zinc finger (FF v).

On reflection, I note that in rejecting claims under 35 U.S.C. § 103, the Examiner bears the initial burden of presenting a *prima facie* case of obviousness. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). In this regard, “[t]he Patent Office has the initial duty of supplying the factual basis for its rejection. It may not, because *it may doubt* that the invention is patentable, resort to speculation, unfounded assumptions or hindsight reconstruction to supply deficiencies in its factual basis.” *In re Warner*, 379 F.2d 1011, 1017, 154 USPQ 173, 178 (CCPA 1967).

Where, as here, the Examiner fails to establish a *prima facie* case, the rejection is improper and will be overturned. *In re Fine*, 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988). Accordingly, I would reverse the rejection of claims 1, 2, 6, 7, and 27-34 under 35 U.S.C. § 103 as being unpatentable over the combination of Greisman and Choo '95.

The majority opinion:

Notwithstanding the foregoing discussion, the majority picks up where the Examiner left off. According to the majority, “Appellants’ elimination of 50% of the amino acids in a specific single position of a zinc finger in a zinc finger polypeptide was ‘the product not of innovation but of ordinary skill and common sense’ . . .” (*supra* 16). This is nonsense. Other than recognizing that work was being performed in this field¹¹, the majority makes no factual finding to support their position.

Nevertheless, to fill in the gaps of their factually deficient reasoning, the majority injects the legal concepts set forth in *Petering* and *Boesch*. These legal concepts fail to fill these gaps.

Petering:

According to the majority, here, “as in *Petering*, one of ordinary skill in the art would ‘envisege each member’ of the genus, including the 8 found in Appellants’ subgenus” (*supra* 14). The majority’s analysis is off base.

In reaching its conclusion, the *Petering* court noted that, while the generic formula in *Petering* was quite broad, “specific preferences” were described in the prior art. *Petering*, at 681, 133 USPQ at 279. Based on those disclosed preferences, the court found that the narrowed generic

¹¹ See *supra* 9: ¶ 8 through 10: ¶ 14; 11: ¶ 22; and 12: ¶ 27 through 13: ¶ 31.

formula essentially disclosed a limited class of approximately twenty compounds. Each was held to have been disclosed by the genus. *Petering*, at 681-682, 133 USPQ at 279-280.

In contrast to *Petering*, the evidence before this panel fails to provide a reasonable expectation that there is a preference for any particular amino acid at position 2 of a second zinc finger or position 6 of a first zinc finger (FF iv, v, xxv, and xxvii). At best, the evidence before this panel suggests that any observed preference reported in the art is preliminary and that before any rule based on an observed preference can be confidently deduced the specificity of individual fingers for DNA triplets must be confirmed (FF xxv). Further, as Greisman expressly states, “there still is no general code that can be used to design optimal zinc finger proteins for any desired target sequence or that can predict the preferred binding site of every zinc finger protein” (FF v).

Contrary, to the majority’s conjecture, the facts on this record do not support a finding that since the prior art zinc finger libraries were created with a limited genus of 16 amino acids at each position, a person of “ordinary skill in the art would ‘envisage *each member*’ of the genus, including the 8 found in Appellants’ subgenus” for the two specific zinc finger positions set forth in Appellants’ claims (*supra* 14). In my opinion, the evidence before this panel would not even suggest the production of a zinc finger polypeptide library wherein every position (or even positions -1, 1, 2, 3, 5, and 6) of the zinc finger polypeptide was limited to those specific 8 amino acids set forth in Appellants’ claimed invention.

Instead, the factual evidence before this panel establishes that that the amino acids in the DNA binding position of a zinc finger protein interrelate

with each other (FF xxiv), they are specific for particular DNA sequences (FF xxvi), and they are constrained by the conformation of the zinc finger itself (FF xxv). The majority's conjecture notwithstanding, those of ordinary skill in the art reached the opposite conclusion (FF iv, v, and xxv-xxvii).

Having failed to clearly state what evidence they believe would have lead a person of ordinary skill in the art to select only those eight amino acids that are set forth in Appellants' claimed invention for position 2 of the second finger or position 6 of the first finger, I am not persuaded by the majority's reliance on *Petering*.

Boesch:

Apparently recognizing the deficiency in their *Petering* analysis, the majority takes a different tack. According to the majority, “[t]he choice of amino acids in zinc fingers is a result-effective variable known to influence their binding to DNA” (*supra* 14). This is true and is, in fact, what all of the references before this panel teach. The problem with this analysis, however, was that nobody knew which amino acids at any given position of a zinc finger could be optimized by routine experimentation to lead to a predictable result. That is why those of ordinary skill in this art randomized the relevant positions of their zinc finger libraries with sixteen amino acids; excluding only those four that they knew did not, or only rarely, appear in the relevant portion of the fingers (FF i).

A conclusory statement is not sufficient to support a finding of obviousness - even under that guise of “routine optimization.” As discussed above, the prior art provides no preferences that would lead to such a

conclusion. To the contrary, as discussed above, the prior art points in the opposite direction – that there are no preferences and therefore no reasonable expectation that a zinc finger polypeptide library could be “optimized” by restricting the number of amino acid choices at position 2 or 6, or any other position.

This concept of the obviousness of discovering optimum or workable ranges by routine experimentation was addressed in *In re Yates*, 663 F.2d 1054, 1056 n. 4, 211 USPQ 1149, 1151 n. 4 (CCPA 1981). In *Yates* the court found that in many instances it is true that it may be obvious to discover optimum or workable ranges by routine experimentation. The *Yates* court noted, however, that the problem

with such “rules of patentability” (and the ever-lengthening list of exceptions which they engender) is that they tend to becloud the ultimate legal issue-obviousness-and exalt the formal exercise of squeezing new factual situations into preestablished pigeonholes. Additionally, the emphasis upon routine experimentation is contrary to the last sentence of section 103.

Stated differently, “it is facts appearing in the record, rather than prior decisions in and of themselves, which must support the legal conclusion of obviousness under 35 U.S.C. § 103.” *In re Cofer*, 354 F.2d 664, 667, 148 USPQ 268, 271 (CCPA 1966); *Ex parte Goldgaber*, 41 USPQ2d 1172, 1176 (BPAI 1995) (“each case under 35 U.S.C. § 103 is decided on its own particular facts.”)

While the majority relies on *Boesch* to support their “optimization” argument (*supra* 15), I note that the holding in *Boesch* was based a critical review of the evidence on that record and specific factual findings that lead the court to conclude that “the prior art would have suggested ‘the kind of

experimentation necessary to achieve the claimed composition . . .” *Boesch*, at 276, 205 USPQ at 219. On this record, the majority has done neither. Accordingly, I am not persuaded by the majority’s unsupported conjecture.

KSR:

Nevertheless, just in case no one buys into their *Petering* or *Boesch* analysis, the majority leaps to *KSR*, asserting that “it’s now apparent ‘obvious to try’ may be an appropriate test in more situations than we previously contemplated” (*supra* 15). According to the majority the reasoning in *KSR* “is applicable here” because,

[w]hen there is motivation

to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp. If this leads to anticipated success, it likely the product not of innovation but of ordinary skill and common sense. In that instance the fact that a combination was obvious to try might show that it was obvious under § 103.

(*supra* 15-16.)

No doubt there was a problem to be solved - every prior art reference before this panel was interested in the same problem - how does one *predictably* design a zinc finger polypeptide to bind a particular DNA sequence. There was, however, *no* identified solution to this problem in the prior art, let alone no “finite number of identified, predictable solutions”. To the contrary, rather than identify predictable solutions for the design of zinc finger polypeptides, the prior art expressly states that none exist (FF iv, v, and xxv). In addition, the prior art emphasizes that the complexity of this

system has frustrated the identification of predictable solutions (FF iv, v, and xxvi-xxvii). Accordingly, it is my opinion that the evidence before this panel fails to support a conclusion that Appellants' invention is obvious because it may have been "obvious to try."

Not to worry, the majority pulls another case out of their hat to address this pesky "predictability" issue. According to the majority,

to the extent selection was "unpredictable," as Appellants argue (Reply Br. 8), unpredictability cannot be equated to nonobviousness when only a finite number of choices are available, as is the case here. *See Pfizer, Inc. v. Apotex, Inc.*, 480 F.3d 1348, 1364, 82 USPQ2d 1321, 1332 (Fed. Cir. 2007) ("obviousness cannot be avoided simply by a showing of some degree of unpredictability in the art so long as there was a reasonable probability of success").

(*Supra* 15.) As I understand the majority's reasoning, since they believe there were only a finite number of choices and a reasonable probability of success there is no unpredictability in this art. For the reasons set forth above, the art does not support this conjecture, and in fact, those of ordinary skill in this art expressly teach the opposite (FF iv, v, and xxv-xxvii).

In addition, in *Pfizer* the Court noted that their conclusion was based on the particular *facts* in that case and on those *facts* the Court found that the prior art provided "ample motivation" to narrow the genus of 53 compounds to a few. *Pfizer*, at 1363-64, 1367, 82 USPQ at 1332-33. On this record, and in contrast to *Pfizer*, the majority makes no particular factual findings to support their conjecture. In this regard, I direct attention to *Takeda Chem. Indus., Ltd. v. Alphapharm Pty., Ltd.*, --- F.3d ---, 83 USPQ2d 1169, 1176-77 (Fed. Cir. 2007), for a discussion of the need to draw conclusions from the facts appearing in the record, rather than prior decisions in and of

themselves, particularly in the context of *KSR* and *Pfizer* as relied upon by the majority on this record.

For the foregoing reasons, I find the majority's conclusion that "[t]he skilled artisan would have had reason to use these methodologies [that were known in the art] to narrow the number of amino acids from 16 to 8 . . ." (*supra* 16) is wholly unsupported by the factual evidence on this record. The same is true of their assertion that "Appellants' elimination of 50% of the amino acids in a single position was 'the product not of innovation but of ordinary skill and common sense . . .' (*id.*).

In all, the majority has built a house of cards on the foundation of nothing more than random excerpts of the prior art and the "per se" application of so-called "rules of patentability." In my opinion, the majority's card house cannot stand up to the weight of the evidence on this record which expressly points in a direction that is opposed to the majority's conclusion. Therefore, contrary to the majority's conjecture, I find that the evidence before this panel is not sufficient to support a conclusion that Appellants' claimed invention would have been *prima facie* obvious over the combination of Greisman and Choo '95, alone or in combination with any of the other references relied upon by the majority.

Claim 30:

Lastly, the majority asserts that the library of claim 30 "is fully disclosed be [sic] Greisman" (*supra* 16). Apparently, the majority is of the opinion that the sequence of eight clones set forth in Fig. 3A of Greisman that were isolated from a library of between 5.6×10^8 and 1.9×10^9 clones,

Appeal 2007-0743
Application 09/424,482

or even a “sublibrary” of this larger library that was screened against the TATA box, is representative of the entire library or sublibrary. Neither Greisman, nor any other reference before this panel, support this assertion. For their part, the majority fails to support this assertion with a direct citation to that portion of Greisman that supports their position. Accordingly, I am not persuaded by their assertion that Greisman anticipates claim 30.

lbg

TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO CA 94111-3834

The opinion in support of the decision being entered today was not written for publication and is not binding precedent of the Board.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Ex parte BRYAN S. WANG, and CARL O. PABO

Appeal No. 2006-3085
Application No. 09/636,243

ON BRIEF



Before ADAMS, MILLS, and LEBOVITZ, Administrative Patent Judges.

LEBOVITZ, Administrative Patent Judge.

DECISION ON APPEAL

This appeal involves claims to zinc finger complexes of two or more fusion proteins, each fusion protein comprising a zinc finger joined to a peptide linker. The Examiner has rejected the claims as lacking written description, indefinite, and anticipated. We have jurisdiction under 35 U.S.C. § 6(b). We reverse the rejections, but set forth a new ground of rejection in which we find all claims obvious in view of prior art.

Background

Zinc finger proteins are DNA-binding proteins. Specification, page 1, lines 11-12. A single finger domain is about 30 amino acids long and recognizes a specific sequence

of nucleotides. Id., page 1, lines 15-16 and 23-30; page 7, lines 18-20. Protein engineering has been utilized to generate zinc fingers with novel DNA sequence specificities. Id., page 2, line 3-page 3, line 16. Because of their DNA-binding properties, zinc finger proteins have been adapted to a variety of different applications, including for sequence detection of target nucleic acid in a sample (id., page 27, lines 6-7) and to regulate gene expression (id., page 27, lines 27-32).

The instant application describes zinc finger proteins that have been appended with peptide sequences which are able to interact and bind to each other. The peptide sequences facilitate the formation of zinc finger protein complexes in which the zinc finger proteins are held together by interactions between the peptides and other contacts the peptides make. Id., page 11, lines 26-34.

Discussion

Claim construction

Claims 5, 6, and 20 are appealed. These are the only pending claims in the application. Claim 5 is the only independent claim and it reads as follows:

5. A zinc finger complex, comprising two or more fusion proteins, each fusion protein comprising a zinc finger protein and a peptide linker, wherein the fusion proteins are joined to each other by specific binding of the peptide linkers, and wherein the peptide linkers are non-naturally occurring peptides.

The claim is drawn to a complex which comprises at least two fusion proteins which are "joined to each other by specific binding of the peptide linkers." To understand this structure, we must first look at the construction of the fusion protein.

A fusion protein, according to the claim, comprises a zinc finger protein and a peptide linker. A "zinc finger protein" is defined in the Specification to be a protein which "binds DNA in a sequence-specific manner." Specification, page 7, lines 18-20. A zinc finger protein can contain from one to thirty-seven individual fingers, each finger of which binds to a defined subsite within the target DNA site. Id., page 14, lines 12-17. A zinc finger protein can be engineered to vary the order and nucleotide sequence specificity of its finger components. Id., page 14, line 22-page 15, line 15; page 7, lines 22-27.

A "peptide linker" is not expressly defined, but in the context of the Specification would be understood by the skilled artisan. The "Summary of the Claimed Invention" refers to both "dimerizing peptides" and "peptide linkers." Id., page 3, lines 22-25 and 26-31, respectively.

The dimerizing peptides "mediate association" of the zinc fingers to which they are attached. Id., page 11, lines 8-14. According to the Specification, a phage display method can be utilized to select the dimerizing peptides. Id., page 11, lines 15-22. In the description of this method, it is stated that "phage display is used for selection of linkers." (Emphasis added.) Id., page 19, line 22. "The method involves the generation of diverse libraries of peptides, typically linked to the same zinc finger protein, followed by affinity selection for phage bearing peptides with dimerizing activity." (Emphasis added.) Id., page 19, lines 22-25. In this context, the skilled artisan would recognize that the peptides which are "linked" to the zinc finger proteins are the "peptide linkers" recited in the claims. When the peptide linkers are selected for their ability to associate with each other (i.e., dimerize), they can also be referred to as "dimerizing peptides."

The limitation in claim 5 that the "fusion proteins are joined to each other by specific binding of the peptide linkers" requires that the peptide linkers attach to each other (i.e., "bind"), the same activity defined for the dimerizing peptides. As the "peptide linkers" have the same activity as the dimerizing peptides and also the same physical association with zinc fingers, we consider the two to be equivalent for the purposes of claim 5.

On page 15 of the Specification, the phrase "peptide linkers" is used to describe peptides which covalently join portions of the zinc fingers together. Not until the peptide linkers are selected for dimerizing activity would they be characterized as dimerizing peptides or the peptide linkers of claim 5 which join the fusion proteins together by "specific binding." Thus, it is clear in the context of the Specification that the claimed peptide linkers are peptides which have been selected for their ability to associate the fusion proteins by specific binding (i.e., homo- or hetero-dimerization as described in the Specification on page 11, lines 8-14).

The peptide linkers are characterized by the claim as "non-naturally occurring peptides." The term "non-naturally occurring" is defined in the Specification to refer to "objects and sequences not found in nature." Id., page 8, lines 3-4. Typical and preferred non-naturally occurring sequences are described. The examples in the Specification describe the selection of peptides with dimerizing activity. See, e.g., id., page 19, lines 24-25. The dimerizing peptides are present in fusion proteins where they are linked to zinc finger proteins. Id., page 19, lines 32-34; page 20, lines 1-7; page 28, lines 15-18. In these cases, the term "peptide" is being used as shorthand for the amino acid sequence which is contained within the fusion protein. It is the amino acid

sequence of the dimerizing peptide which possesses the claimed "specific binding" activity enabling them to dimerize the fusion proteins. Accordingly, we interpret the phrase "non-naturally occurring peptides" to mean that the peptide linker sequence (i.e., the dimerizing "peptide linker") is not found in nature.

Consistent with this construction, the Specification states that the selection of "peptides" having "novel" dimerization motifs makes it less likely that they will react "with natural dimerization interfaces presented by proteins in the cell." Id., page 38, lines 30-34. That is, a naturally occurring sequence (i.e., "peptide linker" of claim 5) within the fusion protein is disfavored because it would be more apt to specifically bind to its natural dimerizing partner within a normal cellular protein than a sequence which is artificial ("non-naturally occurring"). In this context, both the peptide and the "natural dimerization interface" are amino acid sequences.

The fusion proteins are joined "by specific binding of the peptide linkers." The Specification does not define "specific binding" in the context of the peptide linkers, but it does state that the peptides "mediate association" of the fusion proteins and that the "proteins ... bind to each other via the dimerizing peptides." Id., page 11, lines 8-14 and page 13, lines 5-6, respectively. The association/binding is characterized as involving "supporting interactions [that] can include contacts between the peptides and/or contacts between a peptide and another region of the protein." Id., page 11, lines 30-34. It is not a covalent interaction involving covalent bonds.

In sum, the claim is drawn to a complex of at least two fusion proteins, where each protein comprises a zinger finger which is fused to a peptide sequence. The

peptide sequence is not found in nature (i.e., non-naturally occurring). Contacts between the peptide sequences in each fusion protein hold the complex together.

Rejection under § 112, first paragraph for lack of written description

Claims 5, 6, and 20 stand rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter which was not described in the Specification. The Examiner stated that the "as-filed Specification does not describe a claimed zinc finger complex comprising two or more fusion proteins linked by peptide linkers that are non-naturally occurring peptides." Answer, page 4.

"To fulfill the written description requirement, the patent Specification 'must clearly allow persons of ordinary skill in the art to recognize that [the inventor] invented what is claimed.' In re Gosteli, 872 F.2d 1008, 1012, 10 USPQ2d 1614, 1618 (Fed. Cir. 1989). An applicant complies with the written description requirement 'by describing the invention, with all its claimed limitations.' Lockwood v. American Airlines, Inc., 107 F.3d 1565, 1572, 41 USPQ2d 1961, 1966 (1997)." Gentry Gallery v. The Berkline Corp., 45 USPQ2d 1498, 1502-1503 (Fed. Cir. 1998).

Claim 5, as originally filed, reads as follows:

5. A zinc finger complex, comprising a first fusion protein comprising a first zinc finger protein and a first peptide linker and a second fusion protein comprising a second zinc finger protein and a second peptide linker, wherein the first and second fusion proteins are complexed by specific binding of the first and second peptide linkers, and wherein the first and second peptide linkers are nonnaturally occurring peptides.

Comparing amended claim 5, which is now on appeal, to original claim 5, we find almost all its key claimed limitations, including that the fusion proteins each comprising

a "zinc finger protein" and "peptide linker", that the fusion proteins are joined ("complexed" in the original) by "specific binding", and that the peptide linkers are "non-naturally occurring peptides." The only substantial disparity is that amended claim 5 recites that the complex can comprise "two or more fusion protein," while the original did not contain this express limitation. The Examiner argued that the amended claim lacked support in the disclosure, stating that the disclosure "does not describe two linkers each linked to each ZIF [zinc finger protein] wherein the linkers specifically binds to each other to form a fusion of two zinc fingers as schematically depicted by appellants at page 9 of the 8/8/03 Brief." Answer, page 7. As we understand it, the Examiner's position appears to be that there is no written description for complexes which contain more than two zinc finger fusion proteins held together by specifically binding peptide linkers.

To resolve this issue, we turn to the written description. First, we address whether there is support in the Specification for the claimed subject matter of two fusion proteins held together by specific binding between peptide linkers. We discuss this issue in more detail because the Examiner appears to question whether this embodiment was described in the Specification.

As we have construed the claim, the recited "peptide linker" having "specific binding" activity is the same element referred to in the Specification as a "dimerizing peptide." The skilled worker would have gleaned this concept upon reading the Specification in its entirety. For example, in the "Detailed Description," pages 11-14 are devoted to a description of how peptides are selected from peptide libraries, particularly

random peptide libraries, to identify sequences capable of associating two molecules together (i.e., dimerizing).

The relationship between the peptide linkers utilized in the phage library and the dimerization peptides is also shown in Specification Fig. 3. Fig. 3A is a drawing of a DNA construct containing zinc fingers fused to random peptides ("peptide library"). Fig. 3B shows a dimer of fusion protein products of the DNA construct bound to target DNA. It would be evident that the peptide ("peptide library") linked to ZiF12 (a zinc finger protein) depicted in Fig. 3A, after it has been selected for the dimerizing activity, is equivalent to the "dimerizing peptides" of Fig. 3B.

The examples in the Specification also provide support for the written description of the claimed subject matter. On page 32 ("Results"), it is stated that "[t]o select dimerization motifs, we attached random peptides to a DNA-binding domain and selected those fusion proteins that could bind more stably to a symmetric DNA site." The random 15- and 30-amino acid peptides were "expressed at the amino terminus of the first two zinc fingers of Zif268." Id. The zinc finger component mediated binding to target DNA. The binding of monomers – a single fusion protein containing a peptide and zinc finger – and dimers – two of the fusion proteins associated together by the peptide linker – to DNA were characterized. Id., Table 1; page 33, lines 25-page 34, line 7. These experiments clearly demonstrate Appellants' possession of the claimed zinc finger complexes containing two zinger fingers joined to each other by binding between the peptide linkers. We do not find these examples to be incongruent with the Specification's general description as the Examiner contends. Answer, page 9.

We also find support in the Specification for the limitation that the complex contains "two or more" fusion proteins. As pointed out by Appellants (Brief, page 9), it is stated on page 12 of the Specification that dimerizing peptides "are useful for mediating multimerization of zinc finger proteins." A multimer is defined as a "protein made up of more than one peptide chain,"¹ and would include dimers of two fusion proteins (Fig. 3) and trimers of three (Brief, page 9). Although the only examples are of dimers, the Specification refers to these as a "typical application" of the technology, but does not exclude other embodiments. Specification, page 12, lines 21-22. On page 15 of the Specification, it is expressly stated that "[t]wo or more zinc finger proteins can be linked either covalently or by dimerization." Id., page 15, lines 4-5. It is our view this provides explicit support for the claim language (added by amendment on December 27, 2002) "two or more fusion proteins."

For the foregoing reasons, it is our view that the Specification provides an adequate written description of the claimed subject matter. This rejection is reversed.

Rejection under § 112, second paragraph

Claims 5 and 20 stand rejected under 35 U.S.C. § 112, second paragraph, as failing to point out and distinctly claim "the subject matter which the applicant regards as his invention." The Examiner stated that "[i]t is not clear within the claimed context as to what constitutes a [sic] non-naturally-occurring peptide linkers." Answer, page 5. Appellants challenged the rejection, arguing that "non-naturally occurring" is clearly

¹ Zaid et al., Glossary of biotechnology and genetic engineering, FAO Research and Technology Paper, 158 (1999).

defined in the Specification to include only those sequences not found in nature. Brief, page 11.

A specification must conclude with claims "particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention." 35 U.S.C. § 112, second paragraph (2000). The purpose of §112, second paragraph, is to "reasonably apprise those skilled in the art of the scope of the invention." Miles Labs., Inc. v. Shandon, Inc., 997 F.2d 870, 875, 27 USPQ2d 1123, 1126 (Fed. Cir. 1993).

We agree with Appellants that "non-naturally occurring," which is recited in claim 5, would "reasonably apprise" the skilled artisan of the scope of the claimed subject matter. In particular, page 8, lines 3-4, of the Specification states that "non-naturally[] occurring is used to describe objects and sequences not found in nature." We find no merit in the Examiner's arguments characterizing the description of non-naturally occurring in the Specification as inconsistent with the "conventional wisdom of the art," and confusing in its reference to protein folding. Answer, page 12, paragraph 2. The statements referred to by the Examiner described "preferred," "typical," and "some" embodiments, which are examples of non-naturally occurring sequences.

For the reasons set forth above, we reverse this rejection.

Rejection under § 102(b)

Claims 5 and 20 stand rejected under 35 U.S.C. § 102(b) as anticipated by Pomerantz.²

Pomerantz describes the "design of a dimeric zinc finger protein, ZFGD1, containing zinc fingers 1 and 2 from Zif 268 and a portion of the dimerization domain of GAL4." Pomerantz, Abstract. The GAL4 dimerization domain is utilized to associate two chimeric proteins, each containing a zinc finger fused to the GAL4 domain, i.e., a zinc finger-GAL4 fusion. Id., page 966, column 2; Fig. 1. GAL4 is a naturally-occurring protein that contains a coiled-coil dimerization motif that mediates "protein-protein interaction." Id. Binding studies of the fusion protein established that ZFGD1 binds to DNA as a dimer. Id., page 967, column 2.

The Examiner argued that Pomerantz's disclosure of the zinger finger-GAL4 fusion protein anticipates claims 5 and 20.

Pomerantz recites that a portion of Gal4 is used as dimerizing linker. This portion is shown at page 967, col. 1 under the heading section RESULTS i.e., the portion that binds to the 13-residue DNA substite [sic]. Read in the light of the specification definition of a non-naturally occurring peptide linkers e.g., less than 50% (amino acid) with natural sequences the GAL4 (41-100 residues) is less than 50% of the naturally occurring sequence of Gal.

Answer, paragraph spanning pages 13-14, emphasis removed.

Appellants contended that "[t]he claims on appeal clearly require linking of two or more proteins via non-naturally occurring peptides. In contrast, the dimerizing linker used by Pomerantz, namely amino acids 41 to 100 of GAL4, is clearly a naturally-occurring peptide sequence, inasmuch as it is part of the naturally occurring

² Pomerantz et al., (Pomerantz), Biochemistry, 37(4):965-70 (1998).

GAL4 protein." Supplemental Brief, page 8, paragraph 2. We agree with Appellants' position.

Anticipation requires a showing that each element of the claim is identifiable in a single reference. Perricone v. Medicis Pharm. Corp., 432 F.3d 1368, 1375, 77 USPQ2d 1321, 1325 (Fed. Cir. 2005). The GAL4 domain utilized in Pomerantz's fusion protein is a naturally-occurring sequence obtained from the naturally-occurring GAL4 protein. Having interpreted the claims to require that the peptide linker with specific binding ("dimerizing") activity is a non-naturally occurring sequence, we are compelled to conclude that Pomerantz does not teach each and every element of the claimed subject matter. Consequently, we find that the Examiner has failed to establish adequate evidence of prima facie anticipation. This rejection is reversed.

New Grounds of Rejection

Under the provisions of 37 CFR § 41.50(b), we enter the following new grounds of rejection.

Claims 5, 6, and 20 are rejected under 35 U.S.C. § 103(a) as unpatentable over Pomerantz in view of Krylov.³

The Pomerantz publication has been described above for its disclosure of a zinc finger fused to the naturally occurring dimerization domain extracted from the GAL4 protein. Pomerantz's fusion protein differs from the fusion protein contained in the zinc finger complex of claim 5 by having a naturally occurring dimerization domain, instead

³ Krylov et al. (Krylov), The EMBO Journal, 13(12):2849-61 (1994).

of the non-naturally occurring sequence ("peptide linker") required by the claim.

However, in addition to its disclosure of a zinc finger-GAL4 fusion protein, Pomerantz suggests that other dimerizing domains may be appended to the zinc finger.

The dimerization interface also provides opportunities for further elaboration and optimization. As demonstrated by many studies, the coiled-coil interaction motif offers the potential to modify the dimerization domain to increase dimerization affinity or to specifically promote heterodimer formation (see refs 19 and 20 for examples).

....

Dimer contacts of modest affinity may allow self-assembly at the appropriate binding site and thereby reduce the risks of nonspecific (kinetic) trapping that may occur with large covalently linked sets of fingers. Cooperative binding, by giving a more dramatic concentration dependence, may also allow more precise on/off switching in targeted gene regulation.

Pomerantz, page 970, column 1.

A prima facie case of obviousness requires evidence that the prior art disclosed or suggested all of the elements of the claimed invention, and that those skilled in the art would have been motivated to combine those elements with a reasonable expectation of success. See In re Wilson, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970); In re Vaeck, 947 F.2d 488, 493, 20 USPQ2d 1438, 1443 (Fed. Cir. 1991).

Here, Pomerantz does not teach that the zinger finger protein is attached to a non-naturally occurring dimerization domain as required by claim 5, but suggests that such domains be used "to increase dimerization affinity," "to specifically promote heterodimer formation," and "allow more precise on/off switching in targeted gene regulation." Pomerantz, page 970, column 1. This provides the motivation which would have led one of ordinary skill in the art to have replaced the GAL4 dimerization domain with non-naturally occurring sequences having the requisite dimerizing activity.

Pomerantz points the skilled artisan directly to prior art publications that teach modified dimerization domains. Such domains are non-naturally occurring and “join each other by specific binding,” meeting the requirements of the claimed “peptide linkers.” See claim 5. In particular, reference 19 (hereinafter “Krylov”), cited by Pomerantz for its studies of the coiled-coil interaction motif, describes “protein design rules that can be used to modify leucine zipper-containing proteins to possess novel dimerization properties.” Krylov, page 2850, column 1. “33 different leucine zipper proteins containing 27 different systematic combinations of amino acids” were produced. Id., page 2856, column 2 (“Discussion”). See also Fig. 1B for a list of exemplary “mutant proteins.” Id., page 2850, column 2. The mutant proteins were mixed together under conditions which facilitated dimer formation. By measuring the stability of the dimers formed (id., page 2852-53, “Thermodynamic stability”), Krylov was able to demonstrate that certain modified dimers had increased stability and specificity as compared to the unmodified form. (“Novel heterologous interactions regulate dimerization specificity. ... In the second mixing experiment, the stability of the heterodimer is calculated to be greater than the average of the two homodimer stabilities, thus favoring the formation of heterodimers.” Id., page 2856, columns 1-2.) Thus, the element missing from Pomerantz – non-naturally occurring peptide linkers – is supplied by Krylov. The skilled worker would have had a reasonable expectation that Krylov’s domains could be utilized to complex zinc fingers to which they are attached in view of Krylov’s success in not only modifying their binding activity, but in making it stronger (i.e., more stable).

Krylov also teaches dimerization domains having the same sequence, meeting the limitations of claim 6. See e.g., id., page 2856, column 1, describing homo- and heterodimers, where the homodimers have "the same sequence."

Pomerantz describes dimers between ZFGD1 fusion protein, where each fusion contains the same zinc finger. Pomerantz, Abstract ("a dimeric zinc finger protein, ZFGD1"). This meets the requirements of claim 20.

In sum, we find that Pomerantz and Krylov disclose all elements of the subject matter recited in claims 5, 6, and 20. For the reasons discussed above, the skilled worker would have considered these claims obvious in view of Pomerantz's express suggestion to combine its teaching with Krylov (i.e., reference 19), and Krylov's disclosure that would have led the skilled worker to reasonably expect that the combination would work.

Summary

The rejections of claims 5, 6, and 20 are reversed. A new ground of a rejection has been entered under § 103 for claims 5, 6, and 20.

Time Period for Response

This decision contains a new ground of rejection pursuant to 37 CFR § 41.50(b) (effective September 13, 2004, 69 Fed. Reg. 49960 (August 12, 2004), 1286 Off. Gaz. Pat. Office 21 (September 7, 2004)). 37 CFR § 41.50(b) provides "[a] new ground of rejection pursuant to this paragraph shall not be considered final for judicial review."

37 CFR § 41.50(b) also provides that the appellant, WITHIN TWO MONTHS FROM THE DATE OF THE DECISION, must exercise one of the following two options with respect to the new ground of rejection to avoid termination of the appeal as to the rejected claims:

(1) *Reopen prosecution.* Submit an appropriate amendment of the claims so rejected or new evidence relating to the claims so rejected, or both, and have the matter reconsidered by the examiner, in which event the proceeding will be remanded to the examiner. . . .

(2) *Request rehearing.* Request that the proceeding be reheard under § 41.52 by the Board upon the same record. . . .

REVERSED, 37 CFR 41.50(b)



Donald E. Adams
Administrative Patent Judge


Demetra J. Mills
Administrative Patent Judge


Richard M. Lebovitz
Administrative Patent Judge

) BOARD OF PATENT
)
APPEALS AND
)
INTERFERENCES
)

COOLEY GODWARD KRONISH LLP
3000 EL CAMINO REAL
5 PALO ALTO SQUARE
PALO ALTO CA 94306



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/337,216	01/06/2003	Casey C. Case	8325-0009.01	8459
42997	7590	06/30/2008	EXAMINER	
SANGAMO BIOSCIENCES, INC. 501 CANAL BOULEVARD, SUITE A100 RICHMOND, CA 94804			BRUSCA, JOHN S	
			ART UNIT	PAPER NUMBER
			1631	
			MAIL DATE	DELIVERY MODE
			06/30/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Ex parte CASEY C. CASE, EDWARD J. REBAR and
ANDREW JAMIESON

Appeal 2007-2707
Application 10/337,216
Technology Center 1600

Decided: June 30, 2008

Before DONALD E. ADAMS, DEMETRA J. MILLS, and ERIC GRIMES,
Administrative Patent Judges.

MILLS, *Administrative Patent Judge.*

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134. The Examiner has rejected the claims for anticipation and obviousness. We have jurisdiction under 35 U.S.C. § 6(b).

We select claims 31 and 35-39 as representative of the appealed subject matter.

31. A nucleic acid library encoding randomized zinc finger proteins, wherein the library is generated by recognition helix grafting or by gene shuffling.

35. The library of claim 31, wherein each member of the library further encodes a regulatory domain in operative linkage with the zinc finger protein.

36. The library of claim 35, wherein the regulatory domain is an activation domain.

37. The library of claim 35, wherein the regulatory domain is a repression domain.

38. A eukaryotic cell population transfected with the library of claim 31.

39. The cell population of claim 38, wherein the library is generated by recognition helix grafting.

The Examiner relies on the following prior art reference:

Liu et al., *Design of polydactyl zinc-finger proteins for unique addressing within complex genomes*, 94 PROC. NATL. ACAD. SCI. USA 5525-5530 (1997)

Other References of Record

Choo et al. (Choo), *Toward a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage*, 91 PROC.

Appeal 2007-2707
Application 10/337,216

NATL. ACAD. SCI. USA 11163-11167 (1994) (Reference DJ-1 of IDS dated April 4, 2003)

Greisman et al. (Greisman), *A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites*, 275 SCIENCE 657-661 (1997) (Reference EE-1 of IDS dated April 4, 2003)

Jamieson et al. (Jamieson 1), *In Vitro Selection of Zinc Fingers with Altered DNA-Binding Specificity*, 33 BIOCHEMISTRY 5689-5695 (1994) (Reference ET-1 of IDS dated April 4, 2003)

Jamieson et al. (Jamieson 2), *A zinc finger directory for high-affinity DNA recognition*, 93 PROC. NATL. ACAD. SCI. USA 12834-12839 (1996) (Reference ES-1 of IDS dated April 4, 2003)

Rebar et al. (Rebar), *Zinc Finger Phage: Affinity Selection of Fingers with New DNA-Binding Specificities*, 263 SCIENCE 671-673 (1994) (Reference GP-1 of IDS dated April 4, 2003)

Wu et al. (Wu), *Building zinc fingers by selection: Toward a therapeutic application*, 92 PROC. NATL. ACAD. SCI. USA 344-348 (1995) (Reference IB-1 of IDS dated April 4, 2003)

The rejections as presented by the Examiner are as follows:

1. Claims 31, 35, 36 and 38-42 are rejected under 35 U.S.C. § 102(b) as being anticipated by Liu.
2. Claims 31, 35, 37, 38 and 43 are rejected under 35 U.S.C. § 103(a) as being unpatenable over Liu.

We reverse.

DISCUSSION

Background

“The present invention relates to methods of using libraries of randomized zinc finger proteins to identify genes associated with selected phenotypes.” (Spec. 1.) “Zinc finger proteins (‘ZFPs’) are proteins that bind to DNA in a sequence-specific manner and are typically involved in transcription regulation. Zinc finger proteins are widespread in eukaryotic cells.” (Spec. 3.) According to the Specification:

A single finger domain is about 30 amino acids in length and several structural studies have demonstrated that it contains an alpha helix containing the two invariant histidine residues co-ordinated through zinc with the two cysteines of a single beta turn. To date, over 10,000 zinc finger sequences have been identified in several thousand known or putative transcription factors. . . . The structure suggests that each finger interacts independently with DNA over 3 base-pair intervals, with side-chains at positions -1, 2, 3 and 6 on each recognition helix making contacts with respective DNA triplet subsite.

(*Id.* at 3-4.)

“The present application provides for methods of using libraries of randomized zinc finger proteins to screen large numbers of genes, for identifying a gene or genes associated with a selected phenotype” (Spec. 5). In one embodiment, “libraries of randomized zinc finger DNA binding proteins have the ability to regulate gene expression with high efficiency and specificity” (*id.*).

1. Claims 31, 35, 36 and 38-42 are rejected under 35 U.S.C. 102(b) as being anticipated by Liu.

The examiner contends that

Liu et al. shows on page 5528 and throughout construction of two recombinant zinc finger proteins... Liu et al. shows construction of fusions of C7-C7 and Sp1-C7 to the transcriptional activation domain of the herpes simplex virus VP16 protein. In the second column of page 5527 and figure 4 Liu et al. shows human HeLa cells separately transfected with genes that express the two recombinant zinc finger proteins. Liu et al. states in the second column of page 5525 that prior art zinc finger proteins that were designed or selected from libraries of zinc finger proteins have the limitation of comprising zinc finger proteins with only three zinc fingers, which limits the DNA target to a 9 base pair sequence. . . . Liu et al. go on to state in the second column of page 5529 that polydactyl zinc finger proteins will have broad applications for modulation of many different genes in gene therapy applications, and discuss a number of specific gene targets. The discussion of Liu et al. serves to describe a wide variety of recombinant zinc finger proteins.

(Ans. 4-5.) The Examiner further argues that "the claimed subject matter does not require the library of zinc finger protein genes to be in a common container prior to analysis. It is routine in the art to analyze members of a library of genes separately after selection of desired library members, as shown in Liu et al." (Ans. 8).

Appellants contend that "(a) Liu fails to disclose a library, and (b) construction of a library according to claim 31 would never generate only the two polynucleotides disclosed by Liu." (Supp. Br. 6:) Appellants further argue that the "polynucleotides encoding Liu's C7-C7 and Sp1 C-C7 proteins are individual polynucleotides, which Liu does not teach or suggest

can be combined or used together in any way," and, therefore, Liu does not disclose a library. (Supp. Br. 6.) "Appellants reiterate that the two nucleic acids disclosed by Liu are disclosed and analyzed separately, as individual nucleic acids, not as part of any collection that could be considered a library" (Supp. Br. 7). "Appellants believe that the Examiner's interpretation of the term 'library,' to include a collection of only two members, fails to comport with the knowledge of one of skill in the art in the field of molecular biology" (Supp. Br. 8).

Since claim interpretation will normally control the remainder of the decisional process, in considering the issue of patentability "[a]nalysis begins with a key legal question – *what is the invention claimed?*" *Panduit Corp. v. Dennison Mfg. Co.*, 810 F.2d 1561, 1567-1568 (Fed. Cir. 1987). Claim language must be interpreted in light of the claim as a whole, the specification of which the claim is a part, and the prosecution history. See *General Foods Corp. v. Studiengesellschaft Kohle mbH*, 972 F.2d 1272, 1275 (Fed. Cir. 1992). In addition, "[i]t is axiomatic that, in proceedings before the PTO, claims in an application are to be given their broadest reasonable interpretation consistent with the specification and that claim language should be read in light of the specification as it would be interpreted by one of ordinary skill in the art." *In re Sneed*, 710 F.2d 1544, 1548 (Fed. Cir. 1983) (citations omitted). Words in a claim are given their ordinary definition unless they are clearly defined otherwise in the specification. See *Optical Disc Corp. v. Del Mar Avionics*, 208 F.3d 1324, 1334 (Fed. Cir. 2000) ("Without evidence in the patent specification of an express intent to impart a novel meaning to a claim term, the term takes on its ordinary meaning.").

Giving the term "library" its ordinary meaning, we interpret it to mean a pooled collection of different compounds. A cDNA library, for example, is defined as "a pool of complementary DNA clones produced by cDNA cloning of total messenger RNA from a single source (cell type, tissue, embryo)."¹ Appellants further argue that "[l]ibraries, particularly randomized libraries, are constructed for the purpose of generating a large number of different molecules having maximum diversity." (Supp. Br. 8.) Appellants put forth the publications of Choo, Greisman, Jamieson I and II, Liu, Rebar and Wu as further evidence of the meaning of the term "library" to one of ordinary skill in the art, arguing that the references support that libraries are larger than two elements (Supp. Br. 8-9).

In light of the ordinary meaning of the term "cDNA library" as a "pool of DNA clones" and as exemplified by Appellants' evidence we conclude that the term "library" cannot be construed to encompass two individual polynucleotides that are not "pooled" in the same location but are used and analyzed separately as described in Liu.

The standard under § 102 is one of strict identity. "Under 35 U.S.C. § 102, every limitation of a claim must identically appear in a single prior art reference for it to anticipate the claim." *Gechter v. Davidson*, 116 F.3d 1454, 1457 (Fed. Cir. 1997). "Every element of the claimed invention must be literally present, arranged as in the claim." *Richardson v. Suzuki Motor Co., Ltd.*, 868 F.2d 1226, 1236 (Fed. Cir. 1989).

¹(http://www.biochem.northwestern.edu/holmgren/Glossary/Definitions/Def-C/cDNA_library.html).

In view of the above, we find the Examiner has not established a *prima facie* case of anticipation as the examiner has not provided evidence of a library of pooled members, and the rejection of the claims for anticipation over Liu is reversed.

2. Claims 31, 35, 37, 38 and 43 are rejected under 35 U.S.C. 103(a) as being unpatentable over Liu.

The Examiner finds that Liu evidences that

[t]ogether with the data presented here, a general strategy for the generation of proteins with 18-bp recognition sites is the independent evolution of two three-fingered proteins using **phage display** followed by fusion of these proteins with the consensus Kruppel-type linker peptide TGEKP. [Emphasis added.] This strategy should be of general utility in the construction of zinc-finger proteins with genome-specific addressing potential, though exceptions may arise.

(Ans. 6.)

The Examiner finds that Liu further states that “polydactyl zinc finger proteins will have broad applications for modulation of many different genes in gene therapy applications, and discuss a number of specific gene targets. The discussion of Liu et al. serves to describe a wide variety of recombinant zinc finger proteins.” (Ans. 6.)

“Liu et al. shows construction of a fusion of C7-C7 to the transcriptional repression domain of Kruppel associated box-A (KRAB-A) on page 5529 and figure 4. Liu et al. states in the first column of page 5529 that sequence-specific repression may be applicable in gene therapy strategies that target the transcription of viral genomes such as HIV-1. Liu et

al. does not show a plurality of zinc finger proteins fused to a transcriptional repressor." (Ans. 6.)

The Examiner concludes that:

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to construct a second fusion of the Spl-C7 protein to the KRAB-A transcriptional repression domain to further study biological activity of six-finger zinc finger proteins because Liu et al. exemplifies the strategy of assaying activity of zinc fingers fused to transcriptional activation and repression domains to assay for activity of the novel six finger zinc finger proteins constructed in Liu et al. It would be further obvious to construct genes encoding zinc finger proteins fused to transcriptional repression domains because Liu et al. states that repression of gene expression is a useful strategy for repression of viral genome expression. It would have been further obvious to generate a population of cells comprising the recombinant genes to allow for the recombinant genes to be expressed and exert their intended function.

(Ans. 6-7.)

We conclude that the Examiner has not presented a *prima facie* case of obviousness on the evidence before us. In particular, the Examiner has failed to provide evidence of two zinc finger-encoding nucleic acids pooled in a single, pooled location. Nor has the Examiner provided evidence of two pooled zinc finger-encoding nucleic acids or zinc finger-encoding nucleic acids modified to include an activation or repression domain. The Examiner finds that "figure 4 of Liu et al. shows human HeLa cells separately transfected with genes that express the two recombinant zinc finger proteins." (Ans. 5.) The Examiner has cited no evidence to show that the separately transfected cells were pooled.

The Examiner further argues that:

[I]t is known in the art that even pooled populations of cells transfected with a library of genes comprises cells each of which comprises a single member of the library. The appellants have not argued that claim 38 requires a single cell comprising an entire library. The claimed subject matter does not require the population of cells in a single container. Liu et al. shows in figure 4 HeLa cells transfected with the exemplified genes and therefore shows a population of transfected cells.

(Ans. 9.)

We disagree. Consistent with the definition of DNA library set forth herein, we find that the Examiner's evidence fails to show a pooled collection of cells that, individually or collectively, contain multiple zinc finger protein-encoding nucleic acids.

In view of the above, the obviousness rejection is reversed.

SUMMARY

The anticipation rejection is reversed. The obviousness rejection is reversed.

REVERSED

cdc

SANGAMO BIOSCIENCES, INC.
501 CANAL BOULEVARD, SUITE A100
RICHMOND CA 94804

The opinion in support of the decision being entered today was not written for publication and is not binding precedent of the Board.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte ALAN WOLFFE and FYODOR URNOV

MAILED

DEC 21 2006

Appeal No. 2006-2851
Application No. 09/844,501

U.S. PATENT AND TRADEMARK OFFICE
BOARD OF PATENT APPEALS
AND INTERFERENCES

ON BRIEF

Before ADAMS, MILLS, and GRIMES, Administrative Patent Judges.

MILLS, Administrative Patent Judge.

DECISION ON APPEAL

This is a decision on appeal under 35 U.S.C. § 134 from the examiner's final rejection of claims 123-152.

Claims 123 reads as follows:

123. A method for preparing a library of regulatory DNA sequences from a cell, the method comprising:

- (a) providing a cell nucleus, wherein the nucleus comprises cellular chromatin;
- (b) contacting the nucleus with a first enzyme, wherein the first enzyme reacts with accessible regions of cellular chromatin;
- (c) deproteinizing the cellular chromatin to generate deproteinized DNA;

Appeal No. 2006-2851
Application No. 09/844,501

(d) contacting the deproteinized DNA with a second enzyme to generate DNA fragments;

(e) contacting the DNA fragments obtained in step (d) with a population of vector molecules, wherein the vector molecules comprise a first end that is compatible with the first enzyme and a second end that is compatible with the second enzyme, under conditions favorable to ligation of compatible ends; and

(f) selecting polynucleotides comprising a DNA fragment ligated to a vector molecule.

The prior art cited by the examiner is:

Li et al. (Li)	5,500,356	March 19, 1996
Grosveld et al. (Grosveld)	5,635,355	June 3, 1997
Chung	6,644,421	Sept. 3, 2002

NEB Catalog, pp. 32, 46, 48, and 83 (1995)

Grounds of Rejection

Claims 123-128, 130, 135, 143-145, and 147-151 stand rejected under 35 U.S.C. § 103(a) over Grosveld.

Claims 129, 131-133, and 152 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with Li.

Claims 136-142 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with NEB catalog (1995).

Claims 134 and 146 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with Chung.

We reverse these rejections.

Appeal No. 2006-2851
Application No. 09/844,501

DISCUSSION

35 U.S.C. § 103

Claims 123-128, 130, 135, 143-145, and 147-151 stand rejected under 35 U.S.C. § 103(a) over Grosveld.

According to the examiner Grosveld teaches each of the claimed steps, with particular reference to column 8, lines 1-25, column 15, lines 43-47 and column 21, lines 18-20 and claim 1 (Answer, pages 3-4).

Upon review of the disclosure of Grosveld, we do not find the examiner has provided sufficient evidence to support a *prima facie* case of obviousness of the method of claim 123.

We agree with the Examiner that Grosveld describes steps (a)-(d) of the method of claim 123 at Column 8, lines 1-25, we do not find that Grosveld describes a method consistent with steps (e)-(f) of the claimed method.

In particular, claim 123, step (e) recites, "contacting the DNA fragments obtained in step (d) with a population of vector molecules, wherein the vector molecules comprise a first end that is compatible with the first enzyme and a second end that is compatible with the second enzyme, under conditions favorable to ligation of compatible ends."

Grosveld at column 8, lines 16-32, describes deproteination steps and digestion with a second enzyme to generate fragments, such as BgIII, consistent with steps (c)

Appeal No. 2006-2851
Application No. 09/844,501

and (d) of claim 123. Then, the "exact location of the DNasel hypersensitive site[s] of the 3' of the adult β -globin gene were determined using two single copy DNA probes and several restriction enzyme digests of DNasel digested HEL nuclei. The data summarized in FIG. 2 (A-D) show that there is a single DNasel hypersensitive site between the 2.3 kb BglIII fragment and the 2.4 kb HindIII fragment . . ." Column 8, lines 48-51. Accordingly, Grosveld obtained fragments of the adult β -globin gene and probed these fragments to locate the DNasel hypersensitive site. Grosveld did not, according to claim 123, step (e), contact the DNA fragments obtained in step (d) with a population of vector molecules, wherein the vector molecules comprise a first end that is compatible with the first enzyme and a second end that is compatible with the second enzyme, under conditions favorable to ligation of compatible ends; or step (f), select polynucleotides comprising a DNA fragment ligated to a vector molecule. Grosveld, on the other hand, probed DNA fragments which were not ligated to a vector, and selected the DNA fragment of interest having the DNasel hypersensitive site by its ability to bind to a probe.

In a different experiment, Grosveld incorporated the previously identified DNase I hypersensitive sites into a vector or plasmid containing both the hypersensitive sites and the adult β -globin gene. Column 15, lines 6-47. The DNA fragments cloned in the experiment described in column 15 are not the same as the fragments described in column 8. In particular, the hypersensitive site (HSS)-containing fragments cloned in col. 15 are not the DNase/I restriction enzyme fragments from col. 8. See col. 15, lines

Appeal No. 2006-2851
Application No. 09/844,501

45-46: Pvul-BstEII fragment with HSS 1 and 2; BstEII-ClaI fragment with HSS 3 and 4.

In contrast, appellants describe their method in the specification, pages 49-50, as follows.

In another embodiment, cellular chromatin is subjected to limited nuclease action, and fragments having one end defined by nuclease cleavage are preferentially cloned. For example, isolated chromatin or permeabilized nuclei are exposed to low concentrations of DNase I, optionally for short periods of time (e.g., one minute) and/or at reduced temperature (e.g., lower than 37°C). DNase-treated chromatin is then deproteinized and the resulting DNA is digested to completion with a restriction enzyme, preferably one having a four-nucleotide recognition sequence. ... Preferential cloning of nuclease-generated fragments is accomplished by a number of methods. For example, prior to restriction enzyme digestion, nuclease-generated ends can be rendered blunt-ended by appropriate nuclease and/or polymerase treatment (e.g., T4 DNA polymerase plus the 4 dNTPs). Following restriction digestion, fragments are cloned into a vector that has been cleaved to generate a blunt end and an end that is compatible with that produced by the restriction enzyme used to digest the nuclease treated chromatin. ... Ligation of adapter oligonucleotides, to nuclease-generated ends and/or restriction enzyme-generated ends, can also be used to assist in the preferential cloning of fragments containing a nuclease-generated end. For example, a library of accessible sequences is obtained by selective cloning of fragments having one blunt end (corresponding to a site of nuclease action in an accessible region) and one cohesive end ...

In the method of claim 123, it is only after the DNA fragments have been ligated to a vector molecule that the polynucleotide of interest is selected. See, Example 15, specification, page 114, lines 6-14, wherein E.coli colonies harboring insert-containing plasmids were identified and screened.

While both appellants and the examiner rely heavily on argument with respect to potential limitations within the preamble the claims, we do not find it necessary to reach

Appeal No. 2006-2851
Application No. 09/844,501

this issue to decide the case before us. For the reasons discussed herein, we do not find the examiner has provided sufficient evidence to support a *prima facie* case of obviousness. The rejection of the claims over Grosveld is reversed.

35 U.S.C. § 103(a)

Claims 129, 131-133 and 152 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with Li. Claims 136-142 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with NEB catalog (1995). Claims 134 and 146 stand rejected under 35 U.S.C. § 103(a) over Grosveld taken with Chung.

With respect to the other pending obviousness rejections before us, all rejections stand or fall on the relevance of Grosveld to the pending claims. The examiner relies on the NEB catalog to make up for a failure of Grosveld to teach specific restriction enzymes (Answer, page 6), Li for a failure of Grosveld to teach a comparison of cells from a variety of different sources (Answer, page 7), and Chung for the failure of Grosveld to teach embedding cells in agarose prior to enzymatic cleavage (Answer, page 9).

We do not find that either NEB catalog, Li or Chung overcome the above noted deficiency of Grosveld and its failure to teach steps (e) and (f) of claim 123, and therefore the rejections for obviousness over Grosveld taken with NEB catalog, Li or Chung are reversed.

Appeal No. 2006-2851
Application No. 09/844,501

CONCLUSION

The rejections of the claims under 35 U.S.C. § 103(a) over Grosveld alone or in view of NEB, Li or Chung are reversed.

REVERSED



Donald E. Adams
Administrative Patent Judge

)
)
)
)
)
)
)
)


Demetra J. Mills
Administrative Patent Judge

) BOARD OF PATENT
)
) APPEALS AND
)
) INTERFERENCES
)
)


Eric Grimes
Administrative Patent Judge

Appeal No. 2006-2851
Application No. 09/844,501

ROBINS AND PASTERNAK, LLP
1731 Embarcadero Road
Suite 230
Palo Alto, CA 94303



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/083,682	10/24/2001	Alan P. Wolffe	8325-0015.20	1541
20855	7590	09/15/2008		
ROBINS & PASTERNAK			EXAMINER	
1731 EMBARCADERO ROAD			ZHOU, SHUBO	
SUITE 230				
PALO ALTO, CA 94303			ART UNIT	PAPER NUMBER
			1631	
			MAIL DATE	DELIVERY MODE
			09/15/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte ALAN P. WOLFFE and FYODOR URNOV

Appeal 2008-0404
Application 10/083,682
Technology Center 1600

Decided: September 15, 2008

Before DONALD E. ADAMS, DEMETRA J. MILLS, and ERIC GRIMES,
Administrative Patent Judges.

MILLS, *Administrative Patent Judge.*

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134. The Examiner has rejected the claims for anticipation. We have jurisdiction under 35 U.S.C. § 6(b).

STATEMENT OF THE CASE

Claims 66-71 and 125-128 are on appeal. Claims 66 and 67 are representative of the appealed subject matter and read as follows:

66. A polynucleotide, wherein the polynucleotide is a member of a library of polynucleotides, the members of the library comprising a vector and an insert, wherein the insert sequences consist essentially of accessible regions of cellular chromatin, wherein the library is obtained according to the method of:

- (a) contacting cellular chromatin with a probe, wherein reaction of the probe with cellular chromatin results in polynucleotide cleavage at accessible regions of cellular chromatin;
- (b) deproteinizing the cleaved chromatin of step (a);
- (c) digesting the deproteinized chromatin of step (b) with a nuclease to generate a collection of polynucleotide fragments; and
- (d) selectively cloning polynucleotide fragments comprising one end generated by probe cleavage.

67. A library comprising a plurality of polynucleotides according to claim 66.

Cited Reference

Clontech Catalog, (1998-1999) Clontech Laboratories, Inc. 177-183.

Grounds of Rejection

Claims 66-71 and 125-128 stand rejected under 35 U.S.C. § 102(b) as being unpatentable over Clontech.

DISCUSSION

Background

“The present disclosure relates to the fields of bioinformatics, gene regulation, gene regulatory sequences, gene regulatory proteins and methods of determining gene regulatory pathways.” (Spec. 1.)

Claims 66-71 and 125-128 stand rejected under 35 U.S.C. § 102(b) as being unpatentable over Clontech. We select claims 66 and 67 as representative of the rejection before us since Appellants have not separately argued individual claims in the Brief or Reply Brief. 37 C.F.R. 41.37(c)(1)(vii).

The Examiner finds that:

Clontech Catalog discloses multiple genomic libraries made from cellular chromatin of different organisms using different vector systems. See pages 177-183, especially the table on pages 182-183. These genomic libraries are made by a method involving digesting genomic DNA, which is from cellular chromatin, of the different organisms with restriction enzymes, Sau3AI and MboI, which are four cutters and are known in the art to digest the genomes with high frequency, and cloning the digested fragments in different vector systems. See page 177. Given that Sau3AI and MboI are restriction enzymes having recognition sites that occur frequently in the genome, it would be readily apparent to one of skill in the art that the libraries produced by such a method inherently comprise clones that either have an insert that consists entirely of polynucleotide sequence from regions of cellular chromatin that are accessible to reagents such as nuclease and restriction enzymes, as recited in claims 125-128, or have an insert that comprises polynucleotide sequence from the accessible region and sequence from the inaccessible region in the same insert.

(Ans. 4-5.)¹

¹ Reference to the Answer throughout is to the Answer dated May 8, 2007.

The Examiner further finds that:

Due to the use of transitional phrase “consisting essentially of” in claim 66, it is interpreted that the claimed polynucleotide, i.e. the insert, can include sequence from accessible region and sequence other than the accessible region. Thus, the libraries disclosed by Clontech and certain clones contained therein are the same as the polynucleotides or library thereof in the instant product-by-process claims.

(Ans. 5.) The Examiner further argues that the term “consisting essentially of” is not defined in the Specification and thus should be construed as equivalent to the term “comprising.” (*Id.* at 7.)

Appellants contend that the Examiner has misinterpreted the phrase “consisting essentially of” in claim 66 to include “polynucleotides corresponding to inaccessible regions.” (App. Br. 13.)² Appellants argue that “the basic and novel characteristics of the claimed polynucleotides and libraries are that they arise from, and correspond to, accessible regions of cellular chromatin.” (*Id.* at 13.)

Appellants further argue that:

because the claimed libraries are made from cleaving cellular chromatin (in which inaccessible regions are protected from cleavage by associated proteins), the inserts in the library of the claims will never include only inaccessible regions and, moreover if any sequences corresponding to inaccessible regions are present in these inserts, they will be immaterial.

(Reply Br. 3.)³

² Reference to the Appeal Brief throughout is to the Appeal Brief dated December 26, 2006.

³ Reference to the Reply Brief throughout is to the Reply Brief dated July 9, 2007, unless otherwise stated.

We find that when the Specification is viewed as a whole it discloses that the basic and novel characteristics of the claimed polynucleotides and libraries are that they arise from, and correspond to, accessible regions of cellular chromatin. Throughout the Specification Appellants refer to methods of production of libraries which do not include only inaccessible regions of chromatin and to polynucleotides and libraries corresponding to accessible regions of chromatin. (See e.g., Spec. 4:27-32 and Spec. 27: 13-34.) Thus, we do not agree with the Examiner that on the record before us the term "consisting essentially of" should be construed as equivalent to the term "comprising".

Nevertheless, claim 66 is drawn to a polynucleotide. The Examiner finds that, because of the method used to make them, the Clontech libraries inherently "comprise clones that either have an insert that consists entirely of polynucleotide sequence from regions of cellular chromatin that are accessible . . . , or have an insert that comprises polynucleotide sequence from the accessible region and sequence from the inaccessible region in the same insert" (Answer 4-5).

Appellants do not dispute this finding. (Appeal Br. 15: "[D]igestion of naked DNA (as described in Clontech) will necessarily result in a collection (library) of DNA fragments that include both accessible and nonaccessible regions.") Thus, the evidence of record shows that at least one of the polynucleotides of the Clontech libraries will "consist essentially of" accessible regions of chromatin, as claimed. It is of no import that the polynucleotide of accessible regions of chromatin is present in a library

Appeal 2008-0404
Application 10/083,682

because there is no requirement in claim 66 that the polynucleotide be isolated from other polynucleotides.

Appellants argue that “[b]y no stretch of the imagination could digestion of naked DNA, under any circumstances, possibly produce a library of polynucleotides consisting essentially of accessible regions, as claimed.” (App. Br. 14.) We do not find this argument convincing. Again, claim 66 is drawn to a polynucleotide and not to a library. As indicated herein, Clontech discloses a polynucleotide consisting of accessible regions of chromatin, and therefore it is of no consequence that the library of Clontech also includes polynucleotides comprising a mixture of accessible regions and inaccessible regions and polynucleotides comprising accessible regions of chromatin.

Appellants argue that Clontech does not inherently anticipate claims 66 and 67. (Appeal Br. 15.) Because the libraries of Clontech include polynucleotides consisting essentially of accessible regions of chromatin, Clontech inherently discloses the claimed subject matter. As to the library of claim 67, the claimed library “comprises” polynucleotides of accessible regions of cellular chromatin such as those disclosed in Clontech. The term “comprising” is inclusive and does not exclude additional, unrecited elements or method steps. *Moleculon Research Corp. v. CBS, Inc.*, 793 F.2d 1261, 1271 (Fed. Cir. 1986). Because the Clontech libraries comprise polynucleotides of accessible regions of cellular chromatin, Clontech anticipates the library of claim 67.

The anticipation rejection is affirmed.

Appeal 2008-0404
Application 10/083,682

SUMMARY

The anticipation rejection is affirmed.

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a).

AFFIRMED

cdc

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO CA 94303

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte JEFFREY MILLER, GUOFU LI, CARL PABO,
and TREVOR COLLINGWOOD

Appeal 2007-0848
Application 10/651,761
Technology Center 1600

Decided: November 16, 2007

Before DEMETRA J. MILLS, NANCY J. LINCK, and RICHARD M.
LEBOVITZ, *Administrative Patent Judges*.

LINCK, *Administrative Patent Judge*.

DECISION ON APPEAL

This is a 35 U.S.C. § 134 appeal in the above-referenced case.¹

We have jurisdiction under 35 U.S.C. § 6(b). We reverse.

¹ The application was filed August 29, 2003. The real party in interest is Sangamo BioSciences, Inc.

STATEMENT OF THE CASE

The field of the invention is “gene regulation and the production of products” (Specification (“Spec.”) 1). More particularly, Appellants are claiming proteins having multiple engineered zinc finger domains (“multi-ZFPs”) that can bind to and simultaneously regulate several individual genes (Spec. 15-16). According to their Specification, this technique offers a “significant improvement over current technologies by providing the ability to generate a single multiZFP-TF that modulates more than one target gene” (Spec. 16).

The claimed subject matter is reflected in representative claim 1:²

1. (previously presented): A protein comprising two or more engineered zinc finger domains, wherein:
 - (i) each zinc finger domain comprises at least two fingers;
 - (ii) each zinc finger domain binds a different target site; and
 - (iii) the protein modulates expression of two or more endogenous genes.

The Examiner has rejected claims 1-6, 8-10, 12-16, 18, and 19 under 35 U.S.C. § 112, ¶ 1, for lack of written description.

PATENTABILITY UNDER § 112, ¶ 1

The Written Description Issue

Appellants contend they (1) have “amply describe[d] that which is new, i.e., multi-ZFPs to regulate expression of two or more endogenous

² Appellants do not separately argue the claims. Thus, we address the single issue before us with reference to claim 1.

genes” (Reply Br. 15) and (2) do not need to disclose multiple examples of their claimed subject because “methods for designing a zinc finger protein to bind to any target sequence of choice, in any gene, were well-known in the art” (Reply Br. 6).

While admitting the individual components of Appellants’ claimed invention would have been known (Final Office Action (“FOA”) 4), the Examiner finds the skilled artisan would not have known “which zinc finger domain, target site or endogenous gene is involved” in Appellants’ claimed “engineered construct” because Appellants “have not provided an adequate written description” of these features (*id.*).

In view of the above, we frame the written description issue: Does Appellants’ Specification contain a written description sufficient to show they had possession of the claimed constructs, in view of what was known in the art at the time of Appellants’ invention?

Findings of Fact Relating to § 112, ¶ 1

1. Appellants’ claim 1 requires a protein with “two or more engineered zinc finger domains,” each with at least two fingers and each binding a different target site on the protein’s DNA in such a way that “the protein modulates expression of two or more endogenous genes.” (See claim 1.)

2. As admitted by the Examiner, the “prior art describes an ample number of zinc finger domains that bind a variety of target sites” (FOA 4).

3. As further admitted by the Examiner, as a “consequence of the interaction of [these known] zinc finger domains and their targets the expression of endogenous genes is regulated” (FOA 4).
4. Engineered ZFPs are designed “according to the methods and compositions” disclosed in the Specification (Spec. 8) and are well known in the art.
5. Designed ZFPs are proteins “not occurring in nature whose structure and composition results principally from rational criteria” such as “application of substitution rules and computerized algorithms for processing information in a database storing information of existing ZFP designs and binding data” (Spec. 7-8).
6. Selected ZFPs are “protein[s] not found in nature whose production results primarily from an empirical process such as phage display, two-hybrid systems and/or interaction trap assays” which were known in the art (Spec. 8).
7. “[P]owerful strategies for the design of ZFPs . . . that can recognize virtually any desired DNA sequence . . . have been developed” (Spec. 14-15 (citing numerous references)).
8. “Selection methods for obtaining, from randomized libraries, zinc fingers that bind any particular target sequence of choice are described” in the prior art (Reply Br. 5-6; *see also* references cited at Reply Br. 6).
9. Methods for “designing a zinc finger protein to bind to any target sequence of choice, in any gene, were well-known in the art (*id.*).

10. The Specification discloses in detail a working example involving the design and selection of a multiZFP-TF designed to modulate gene expression in tocopherol (Spec. 18-25).

11. Given the advanced state of the zinc finger art (as evidenced by the multiple references cited in the Specification), and Appellants' detailed teachings regarding modulating gene expression in tocopherol (Spec. 18-25), the skilled artisan would have concluded Appellants were in possession of the claimed invention (FF 2-10).

Discussion of the Written Description Issue

Appellants claim a genus of proteins that "modulate[] expression of two or more endogenous genes" and comprise "two or more engineered zinc finger domains" (claim 1). There is no dispute that much is known in the zinc finger art (FF 2-10), a fact evidenced by the many, many references cited by Appellants in their Specification and as an exhibit to their Reply Brief. Thus, the only question before us is whether Appellants' combination, requiring modulation of the expression of more than one endogenous gene by more than one engineered zinc finger domain was sufficiently described to show possession.

The law on written description is less than clear, as is evidenced by the caselaw relied upon by the Examiner and Appellants. However, we agree with Appellants that recently decided *Capon v. Escher*, 418 F.3d 1349 (Fed. Cir. 2005) is more applicable to the facts of this case than is *Regents v. Eli Lilly & Co.*, 119 F.3d 1559 (Fed. Cir. 1997). In *Capon*, the court noted

what is needed to support generic claims to biological subject matter [such as are before us] depends on a variety of factors, such as the existing knowledge in the particular field, the

extent and content of the prior art, the maturity of the science of technology, the predictability of the aspect at issue, and other considerations appropriate to the subject matter.

418 F.3d at 1359.

In this case, analysis of these factors supports Appellants' position: Existing knowledge regarding the design and use of zinc fingers is extensive; the relevant prior art is voluminous; zinc finger technology is relatively mature; and based on existing knowledge and the prior art teachings, the outcome of Appellants' work is relatively predictable. (*See* FF 1-10.) Thus, the skilled artisan would have concluded Appellants were in possession of the claimed invention (FF 11). Accordingly, we reverse the Examiner's written description rejection of all the pending claims.

CONCLUSION

In summary, we reverse the written description rejection under 35 U.S.C. § 112, ¶ 1, of claims 1-6, 8-10, 12-16, 18, and 19.

REVERSED

Ssc

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO, CA 94303

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte
H. STEVE ZHANG and LEI ZHANG

Appeal 2008-2455
Application 11/101,095
Technology Center 1600

Decided: September 29, 2008

Before TONI R. SCHEINER, DONALD E. ADAMS, and LORA M. GREEN,
Administrative Patent Judges.

SCHEINER, *Administrative Patent Judge.*

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134 involving claims 1-3 and 6, directed to a method for increasing cardiac contractility in a subject. The Examiner has rejected the claims as lacking enablement, and as obvious over the prior art. We have jurisdiction under 35 U.S.C. § 6(b).

We reverse.

BACKGROUND

“Congestive heart failure is a syndrome characterized by left ventricular dysfunction . . . Decreased contractility of the left ventricle leads to reduced cardiac output with consequent systemic arterial and venous vasoconstriction” (Spec. 1: 12-15). “Contractility appears to be regulated primarily by calcium flow” (Spec. 1: 16).

“Phospholamban (PLN) is a regulatory phosphoprotein that modulates the active transport of Ca²⁺ by the cardiac sarcoplasmic reticular Ca(2+)-ATPase enzyme (SERCA2) into the lumen of the sarcoplasmic reticulum” (Spec. 1: 25-28).

The Specification describes several zinc finger proteins (ZFPs) that bind to target sites on the phospholamban gene, as well as fusion proteins wherein the ZFPs are fused to a transcriptional repression domain, e.g., a KOX domain (Spec. 2: 20-24, 3: 25-26).

According to the Specification, “repression of PLN expression can be achieved using the ZFP[] [fusion proteins] described herein, thereby increasing contractility (e.g., by increasing SERCA2a:PLN ratio). Thus, . . . the ZFPs can be used to repress expression of PLN, both *in vitro* and *in vivo*. Such repression can be utilized for example to alter the contractile activity of cardiac muscle and, accordingly, as treatment for congestive heart failure” (Spec. 20: 17-21).

The present invention is directed to a method of increasing cardiac contractility in a subject by administering a fusion protein comprising a phospholamban-targeted ZFP and a repression domain.

DISCUSSION

CLAIMS

Claim 1 is representative and reads as follows:

1. A method for increasing cardiac contractility in a subject, the method comprising:

introducing a nucleic acid into the subject, wherein the nucleic acid encodes a polypeptide, wherein the polypeptide comprises:

(i) a zinc finger DNA-binding domain that is engineered to bind to a target site in the phospholamban gene; and

(ii) a transcriptional repression domain;

such that the nucleic acid is expressed in one or more cardiac cells of the subject, whereby the polypeptide binds to the target site and represses transcription of the phospholamban gene.

Claims 1-3 and 6 stand rejected under 35 U.S.C. § 112, first paragraph, as lacking enablement throughout their scope.

Claims 1 and 6 stand rejected under 35 U.S.C. § 103(a) as unpatentable over Choo,¹ GenBank,² Rebar,³ and del Monte.^{4, 5}

ENABLEMENT

The Examiner rejected claims 1-3 and 6 under 35 U.S.C. § 112, first paragraph, as nonenabled, on the basis that the “nature of the claimed

¹ International Application WO 01/25417 A2 of Choo, published April 12, 2001.

² GenBank Accession No. AF177763 (dated September 21, 1999) (hereinafter “GenBank”).

³ International Application WO 02/057293 A2 of Rebar, published July 25, 2002.

⁴ F. del Monte et al., *Targeting Phospholamban by Gene Transfer in Human Heart Failure*, 105 Circulation 904-7 (2002).

⁵ This ground of rejection has been withdrawn by the Examiner with respect to claims 2-5 (Ans. 29).

invention can be reasonably construed as a gene therapy method" (Ans. 7), "for use in treating heart failure" (*id.*). "However, [the] specification fails to demonstrate . . . modulation of any global cardiac function indicating that exemplified gene transfer would result in . . . functionally meaningful expression for [a] sustained period to achieve any therapeutic response" (Ans. 12).

The Examiner acknowledges that "claims 1-3 and 6 are drawn to a method for increasing cardiac contractility" in a subject (Ans. 5), and that "[t]he specification provide[s] guidance with respect to administering plasmids encoding 6439-KOX . . . directly into adult rat myocardium" and "[t]he data shows 6439-KOX increases Ca²⁺ transients in isolated adult rat cardiomyocytes" "isolated after injection" (Ans. 10). Nevertheless, the Examiner contends that "the claimed invention can be reasonably construed as a gene therapy method" (Ans. 7), thus, the claims have been analyzed accordingly, and "have been also analyzed for their intended use in the treatment of heart failure and other cardiac disorders" (Ans. 5).

In this regard, the Examiner cites a number of references⁶ in support of the assertion that "[t]he state of the art of gene therapy or delivering nucleic acids at the time of the filing of this application was unpredictable wherein any gene was expressed in an individual" (Ans. 9).

Appellants argue "the claims are not drawn to treatment methods or to methods of maintaining a sustained pharmacological response" (Reply Br.

⁶ For example, the Examiner cites Verma & Somia, *Gene Therapy - Promises, problems and Prospects*, 389 Nature 239-242 (1997), and Pfeifer & Verma, *Gene Therapy: Promises and Problems*, 2 Annu. Rev. Genomics Hum. Genet. 177-211 (2001).

4), thus the claims do not “necessitate any showing . . . [of] *in vivo* treatment of heart failure resulting from a sustained ‘pharmacological response.’ All that is required is that the specification teach one of skill in the art how to increase cardiac contractility using PLN-repressing ZFPs” (*id.*). Appellants contend that “[t]he as-filed specification includes working examples showing enablement of the claimed methods both *in vitro* (repression of PLN) and *in vivo* (repression of PLN increases cardiac contractility)” (*id.*). Thus, Appellants argue, “the working examples exemplify and adequately support the claimed methods” (*id.*).

We agree with Appellants that the Examiner has applied an overly stringent standard for enablement in this case. “[T]o be enabling, the specification of a patent must teach those skilled in the art how to make and use the full scope of the claimed invention without ‘undue experimentation.’” *In re Wright*, 999 F.2d 1557, 1561 (Fed. Cir. 1993). “That *some* experimentation may be required is not fatal; the issue is whether the amount of experimentation required is ‘undue.’” *In re Vaeck*, 947 F.2d 488, 495 (Fed. Cir. 1991) (emphasis in original). Some experimentation, even a considerable amount, is not “undue” if, for example, the specification provides a reasonable amount of guidance as to the direction in which the experimentation should proceed. *In re Wands*, 858 F.2d 731, 737 (Fed. Cir. 1988).

The enablement analysis must be focused on the product or method defined by the claims. “Title 35 does not require that a patent disclosure enable one of ordinary skill in the art to make and use a perfected, commercially viable embodiment absent a claim limitation to that effect.”

CFMT, Inc. v. Yieldup Int'l Corp., 349 F.3d 1333, 1338 (Fed. Cir. 2003). See also *In re Cortright*, 165 F.3d 1353, 1359 (Fed. Cir. 1999) (claims to method of “restoring hair growth” encompassed achieving full head of hair but did not require it).

The claims on appeal are directed to a method of increasing cardiac contractility in a subject, not to a method of treating heart failure. It is true that the Specification contemplates the use of the claimed method to treat heart failure and other cardiac disorders, but practicing the claimed method does not require a therapeutically effective result.

Moreover, the Examiner’s apparent position that the Specification cannot teach how to use the claimed method unless it teaches solutions to all the problems in the field of gene therapy is contrary to controlling case law. See, e.g., *In re Brana*, 51 F.3d 1560, 1568 (Fed. Cir. 1995).

In *Brana*, the claims were directed to compounds disclosed as anti-cancer agents. *Id.* at 1562. The USPTO rejected the claims as nonenabled, *id.* at 1563-64, despite working examples in Brana’s specification showing treatment of cancer in a mouse model. *Id.* at 1562-63. The USPTO argued that the results of the mouse testing “are not reasonably predictive of the success of the claimed compounds for treating cancer in humans.” *Id.* at 1567. The court concluded that this position “confuses the requirements under the law for obtaining a patent with the requirements for obtaining government approval to market a particular drug for human consumption.” *Id.* The *Brana* court held that “[u]sefulness in patent law, and in particular in the context of pharmaceutical inventions, necessarily includes the expectation of further research and development. The stage at which an

invention in this field becomes useful is well before it is ready to be administered to humans." *Id.* at 1568.

Here, the claims are simply directed to a method for increasing cardiac contractility in a subject, and Appellants' Specification provides several working examples demonstrating just that in rats. The Examiner has interpreted the claims as being directed to a method of treating heart failure, and has discounted the Specification's working examples because they do not demonstrate "express[ing] ZFP-KOX *in vivo* for [a] sustained period showing that [the] instant method could be achieved *in vivo* in substantial number[s] of cardiac cells in order to elicit any pharmacological response" (Ans. 9). However, enablement - especially in the context of pharmaceutical inventions - includes an expectation of further research and development. In the pharmaceutical field, an invention can be enabled well before it is ready to be administered to humans. Thus, enablement is not precluded even if the claims encompass methods, such as gene therapy, that have not yet overcome all the obstacles to their clinical use.

The Examiner has not established that undue experimentation would have been required to practice the *claimed* method; specifically, a method of increasing cardiac contractility in a subject. The claims do not require therapeutically effective treatment of any disease, and we do not concur with the Examiner's conclusion that such an effect is required to satisfy 35 U.S.C. § 112, first paragraph.

The rejection of claims 1-3 and 6 for lack of enablement under the first paragraph of 35 U.S.C. § 112 is reversed.

OBVIOUSNESS

The Examiner rejected claims 1 and 6 under 35 U.S.C. § 103(a) as unpatentable over Choo, GenBank, Rebar, and del Monte.

Choo describes a library of DNA sequences immobilized on a solid support that “allows improved selection for zinc fingers with particular sequence identity” (Choo 5: 20-21). Choo does not disclose ZFPs specific for phospholamdan.

Rebar describes a fusion protein “comprising a modified zinc finger DNA-binding domain and a functional domain . . . for . . . [repressing] endogenous gene expression” (Rebar 21: 1-3), for example, a ZFP DNA-binding domain fused to a KRAB repression domain from the human KOX-1 protein (Rebar 21: 9-11). Rebar does not disclose fusion proteins specific for phospholamdan.

GenBank discloses the sequence of the phospholamdan genes for several species.

del Monte teaches that decreasing the level of phospholamdan using an antisense approach improves contractility in failing cardiomyocytes (del Monte Abstract).

The Examiner contends that “[i]t would have been obvious for one of ordinary skill in the art . . . to modify the method of del Monte by replacing . . . anti-sense-PLN . . . with a nucleic acid encoding polypeptide comprising zinc finger DNA-binding fusion protein specifically designed to target [PLN]” (Ans. 16-17), because “GenBank had already disclosed the . . . sequence of different regions of PLN in different species” (*id.* at 17), while Choo “had disclosed method to design a composition to target a target gene”

(*id.*), and Rebar “had shown that a transcription repressor domain could be fused to ZFP-binding domain comprising plurality of zinc fingers for the inhibition of expression of the target gene” (*id.*). The Examiner contends that “[o]ne who would practice the invention would have had reasonable expectation of success because del Monte et al had already described decreasing phospholamdan expression restores contractility in failing ventricular cells of heart” (*id.*).

Appellants contend that “the cited references do not teach all the elements of appealed claims 1 and 6” (Reply Br. 11). Specifically, “del Monte and GenBank . . . are silent as to ZFPs entirely and Choo does not teach nucleic acids encoding a fusion of a ZFP and a repression domain, let alone a PLN-specific ZFP fused to repression domain as claimed. Rebar also fails to teach or suggest PLN-repressing ZFPs” (*id.*). Appellants contend that “[n]one of these references provide any suggestion of increasing cardiac contractibility by repressing PLN expression” rather than destroying phospholamdan mRNA, and moreover, “del Monte’s antisense RNA has not been shown to be interchangeable with ZFP technology” (App. Br. 15).

Appellants have the better argument. Given the complete lack of any teaching regarding PLN-specific ZFPs, the Examiner has not established that one of skill in the art would have had a reasonable expectation of success in essentially “starting from scratch,” and therefore the Examiner has not established a *prima facie* case of obviousness for the claimed invention.

Accordingly, the rejection of claims 1 and 6 as unpatentable over the prior art is reversed.

Appeal 2008-2455
Application 11/101,095

SUMMARY

The rejection of claims 1-3 and 6 under 35 U.S.C. § 112, first paragraph, as lacking enablement is reversed. The rejection of claims 1 and 6 under 35 U.S.C. § 103(a) is reversed as well.

REVERSED

lp

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO, CA 94303

8325-2001.10

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte YEH CHOO, AARON KLUG,
and ISIDRO SANCHEZ-GARCIA

Appeal 2009-010838
Application 10/397,930
Technology Center 2009-010838

Decided: August 28, 2009

RECEIVED

SEP - 2 2009

ROBINS & PASTERNAK LLP

Before TONI R. SCHEINER, ERIC GRIMES, and RICHARD M.
LEBOVITZ, *Administrative Patent Judges.*

GRIMES, *Administrative Patent Judge.*

DOCKETED D

DECISION ON APPEAL

This is an appeal under 35 U.S.C. § 134 involving claims to fusion proteins that include a non-natural zinc finger domain. The Examiner has rejected the claims as obvious. We have jurisdiction under 35 U.S.C. § 6(b). We reverse.

STATEMENT OF THE CASE

Claims 42-44 and 52 are on appeal.¹ Claim 42, the only independent claim, reads as follows:

42. An isolated polypeptide comprising a non-naturally occurring zinc finger protein and a heterologous functional domain,

(a) wherein the non-naturally occurring zinc finger protein:

(i) has been designed to bind to a particular chromosomal target DNA sequence, and

(ii) comprises a non-naturally occurring zinc finger comprising an antiparallel beta sheet packed against an alpha helix; and

(b) wherein the heterologous functional domain is selected from the group consisting of an activation domain, a repression domain, a catalytic domain from a restriction enzyme and a nuclear localization signal.

The claims stand rejected under 35 U.S.C. § 103(a) as follows:

- Claims 42, 43, and 52 in view of Desjarlais,² Courey,³ Anderson,⁴ and Narayan⁵ (Answer 4); and

¹ Claims 45-47 and 50 are also pending but have been withdrawn from consideration by the Examiner (Office action mailed Oct. 15, 2008, p. 1).

² Desjarlais et al., "Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins," 90 Proc. Natl. Acad. Sci. USA 2256-2260 (1993).

³ Courey et al., "Synergistic Activation by the Glutamine-Rich Domains of Human Transcription Factor Sp1," 59 Cell 827-836 (1989).

⁴ Anderson et al., "Synergistic Activation of a Human Promoter in vivo by Transcription Factor Sp1," 11 Molecular and Cellular Biology 1935-1943 (1991).

⁵ Narayan et al., "Structures of Zinc Finger Domains from Transcription Factor Sp1," 272 Journal of Biological Chemistry 7801-7809 (1997).

- Claim 44 in view of Desjarlais, Courey, Anderson, Narayan, and Sadowski⁶ (Answer 7-8).

OBVIOUSNESS

Issue

The Examiner has rejected claims 42, 43, and 52 in view of Desjarlais, Courey, Anderson, and Narayan. The Examiner finds that Desjarlais discloses a non-naturally occurring zinc finger protein that was designed to bind to a particular target DNA (Answer 4-5) and that Narayan shows that Desjarlais' protein comprises the beta sheet/alpha helix structure recited in the claims (*id.* at 5).

The Examiner also finds that Courey discloses a zinc-finger protein comprising a heterologous activation domain (*id.* at 5-6) and that Anderson discloses that a zinc-finger protein binds to binding sites that have been inserted into chromosomal DNA (*id.* at 6). The Examiner concludes that it would have been obvious to modify Desjarlais' protein to include a heterologous activation domain, as taught by Courey, and to insert the binding site of the fusion protein into a chromosome, as taught by Anderson, "in order to receive the expected benefit of making a protein that is capable of activating transcription" and "put the target site into a cellular context which closely resembles that of a single-copy gene" (*id.* at 7).

Appellants contend that "[t]here is nothing in the references or art as a whole that would suggest that it was predictable that proteins as disclosed in Desjarlais (the only reference . . . cited for teaching a non-naturally

⁶ Sadowski et al., "GAL4-VP16 is an unusually potent transcriptional activator," 335 Nature 563-564 (1988).

occurring zinc finger protein, although not one that is designed to bind a chromosomal target site) would bind to chromosomal targets" (Appeal Br. 8). Appellants also contend that the claims "clearly require that the non-naturally occurring zinc finger protein be designed to bind a chromosomal target and . . . it is clear that the target is within a chromosome when it is bound by the non-naturally occurring protein" (Reply Br. 7).

In response to Appellants' argument that the references do not teach a fusion protein designed to bind a chromosomal target DNA sequence, the Examiner reasons that a

chromosomal target DNA sequence is a target DNA sequence found in a chromosome. A sequence synthesized *in vitro* could be the same sequence found in a chromosome. Furthermore, Anderson et al teach it is within the skill of the art to insert a DNA target sequence into a chromosome. . . . Thus, any DNA sequence may be described as a chromosomal target sequence.

(Answer 11-12.)

The issue presented is: Did the Examiner err in concluding that the cited references would have suggested a fusion protein "designed to bind to a particular chromosomal target DNA sequence," as required by the claims?

Findings of Fact

1. Desjarlais discloses three zinc-finger proteins designed to bind to three different DNA sequences, each nine nucleotides in length (Desjarlais, abstract).
2. Desjarlais assayed binding of the proteins to their predicted binding sites using purified proteins (*id.* at 2256, right col.); i.e., *in vitro*.

3. Courey discloses a fusion protein comprising the naturally occurring zinc-finger domain of transcription factor Sp1 and a glutamine-rich domain of the Drosophila Antennapedia protein (Courey 828, left col.).

4. Courey discloses that the fusion protein stimulated expression in cells of “a reporter plasmid containing the CAT gene under control of the Sp1-dependent SV40 early promoter” (*id.*).

5. Anderson discloses that three Sp1 binding sites in the human arginosuccinate synthetase (AS) promoter “were mutated to abolish Sp1 binding, individually and in all possible combinations, to generate a series of AS promoter-chloramphenicol acetyltransferase (CAT) expression constructs” (Anderson, abstract).

6. Anderson discloses that plasmids comprising the mutant AS promoter-CAT constructs were stably transfected into a human cell line (*id.* at 1939, paragraph bridging the columns).

7. Stably transfected cell lines have the exogenous DNA integrated into the chromosome of the cells. (*See id.*: “Any positional effect of the chromosomal integration site on expression of the transfected plasmids should be averaged out.”)

8. The Specification discloses “a method of altering the expression of a gene of interest in a target cell, comprising: determining (if necessary) at least part of the DNA sequence of the structural region and/or a regulatory region of the gene of interest; designing a zinc finger polypeptide to bind to the DNA of known sequence, and causing said zinc finger polypeptide to be present in the target cell” (Spec., col. 7, l. 64 to col. 8, l. 4).

9. The Specification provides a working example describing “a three finger polypeptide able to bind site-specifically to a unique 9 bp region of the BCR-ABL fusion oncogene and to discriminate it from the parent genomic sequences. . . . Using transformed cells in culture as a model, it is shown that binding to the target oncogene in chromosomal DNA is possible, resulting in blockage of transcription” (*id.* at col. 25, ll. 53-60).

Principles of Law

“[T]he PTO applies to the verbiage of the proposed claims the broadest reasonable meaning of the words in their ordinary usage as they would be understood by one of ordinary skill in the art, taking into account whatever enlightenment by way of definitions or otherwise that may be afforded by the written description contained in the applicant’s specification.” *In re Morris*, 127 F.3d 1048, 1054 (Fed. Cir. 1997).

“A claim can be obvious even where all of the claimed features are not found in specific prior art references, where ‘there is a showing of a suggestion or motivation to modify the teachings of [the prior art] to the claimed invention.’” *Ormco Corp. v. Align Technology Inc.*, 463 F.3d 1299, 1307 (Fed. Cir. 2006).

However, 35 U.S.C. § 103 requires “a searching comparison of the claimed invention – including all its limitations – with the teaching of the prior art.” *In re Ochiai*, 71 F.3d 1565, 1572 (Fed. Cir. 1995).

“In rejecting claims under 35 U.S.C. § 103, the examiner bears the initial burden of presenting a *prima facie* case of obviousness.” *In re Rijckaert*, 9 F.3d 1531, 1532 (Fed. Cir. 1993).

Analysis

The claims are directed to a fusion protein comprising a non-naturally occurring zinc finger protein and a heterologous functional domain, where the zinc finger protein “has been designed to bind to a particular chromosomal target DNA sequence” (claim 42). The Examiner acknowledges that none of the references relied on to show obviousness disclose a zinc finger protein that binds to a chromosomal DNA target, but reasons that “chromosomal DNA” includes any DNA that is, or could be, part of a chromosome (Answer 11-12), and since Anderson teaches how to insert exogenous DNA into a chromosome, “any DNA sequence may be described as a chromosomal target sequence” (*id.* at 12).

We disagree with the Examiner’s interpretation of the claim language. To those of ordinary skill in the art, “chromosomal DNA” denotes a category of DNA that occurs naturally as part of a chromosome, as opposed to synthetic DNA having a sequence determined by the synthesizer. The Specification also makes clear that a chromosomal target sequence is one that resides in a chromosome when it is bound by a zinc finger protein. See FF 9 (“a three finger polypeptide able to bind site-specifically to a unique 9 bp region of the BCR-ABL fusion oncogene . . . show[s] that *binding to the target oncogene in chromosomal DNA* is possible,” emphasis added). In addition to this, the Specification refers to designing zinc finger proteins to other natural occurring mutations that reside in the genomic, and therefore chromosomal, DNA (col. 32, l. 10 to col. 32, 56). Therefore, when read in light of the Specification, “a particular chromosomal target DNA sequence”

means a particular sequence that is found in a chromosome as a naturally occurring part of it.

The Examiner has not shown that a zinc finger fusion protein that binds to a particular sequence that is found in a chromosome as a naturally occurring part of it is either disclosed or suggested by the references relied on to show obviousness. Because the Examiner has not shown that a product meeting all of the limitations of the claims would have been obvious based on the prior art and the knowledge of a person of ordinary skill, she has not made out a *prima facie* case of obviousness under 35 U.S.C. § 103.

Conclusions of Law

The Examiner erred in concluding that the cited references would have suggested a fusion protein “designed to bind to a particular chromosomal target DNA sequence,” as required by the claims.

SUMMARY

We reverse the rejection of claims 42, 43, and 52 based on Desjarlais, Courey, Anderson, and Narayan and the rejection of claim 44 based on Desjarlais, Courey, Anderson, Narayan, and Sadowski.

REVERSED

dm

ROBINS & PASTERNAK
1731 EMBARCADERO ROAD
SUITE 230
PALO ALTO, CA 94303